Volatility forecasting in the Chinese commodity futures market with intraday data

Ying Jiang* Shamim Ahmed† Xiaoquan Liu‡

December, 2015

Abstract

Given the unique institutional regulations in the Chinese commodity futures market as well as the characteristics of the data it generates, we utilize contracts with three months to delivery, the most liquid contract series, to systematically explore volatility forecasting for Aluminum, Copper, Fuel Oil, and Sugar at the daily and three intraday sampling frequencies. We adopt popular volatility models in the literature and assess the forecasts obtained via these models against alternative proxies for the true volatility. Our results suggest that the long memory property is an essential feature in the commodity futures volatility dynamics and that the ARFIMA model consistently produces the best forecasts or forecasts not inferior to the best in statistical terms.

JEL Classification: C5, G12, G13.

Keywords: Out-of-Sample Predictability, Long Memory Time Series, Futures Market Regulation, Realized Volatility, and Econometric Models.

*Nottingham University Business School China, University of Nottingham Ningbo, Ningbo 315100, China. E-mail: ying.jiang@nottingham.edu.cn.

†Nottingham University Business School, University of Nottingham, Nottingham NG8 1BB, United Kingdom. E-mail: shamim.ahmed@nottingham.ac.uk.

‡Corresponding author. Nottingham University Business School China, University of Nottingham Ningbo, Ningbo 315100, China. E-mail: xiaoquan.liu@nottingham.edu.cn. Phone: +86 574 88180207. Fax: +86 574 88180205.

1 Introduction

In this paper, we are concerned with volatility forecasting in the Chinese commodity futures market. Volatility modeling and forecasting is a much devoted area of research as volatility is considered the "barometer for the vulnerability of financial markets and the economy" (Poon and Granger (2003, p.479)) and central to asset pricing, derivative valuation, portfolio allocation, and risk management. We are interested in this particular market in part because it has become an important part of the global futures markets with tremendous trading volume.[1,2] More importantly, this market is regulated by two unique institutional rules that makes it interesting to explore.

The first regulation is the time-dependent margin rate, whereby the margin as a fraction of the contract value increases as contracts move closer to delivery. Take Sugar as an example. The margin rate for deposit two months prior to delivery is 6% of the contract value for an investor. In the month before delivery, it increases to 8% in the first 10 days, to 15% between the 11th to the 20th day of the month, to 25% in the final 10 days of the month, culminating to 30% in the delivery month.[3] The second regulation is that, although they represent 97% of all investors in the futures markets, individual investors are not allowed to trade nearby contracts.[4] Both regulations effectively push market participation and trading volume to more distant contracts with implications for market liquidity.

Our contribution to the literature is that we take into account of unique institutional regulations of this market and design empirical volatility forecasting exercises that are appropriate for the characteristics of the market and the data it generates. Our data on Aluminum, Copper, and Fuel Oil consistently show that contracts with three months to delivery enjoy the best liquidity. We are not the first to note this pattern (see Liu et al. (2014) and Peck (2008)), but

[1] See the *Annual Volume Survey Report 2014* published by the Futures Industry Association, the primary industry association for centrally cleared futures and swaps based in Washington D.C., at https://fia.org. The Chinese Sugar futures contracts rank 3rd globally in terms of trading volume in the Agricultural Category, while Copper ranks 4th in the Metals Category.

[2] Our paper is related to Liu et al. (2014) which examine hedging with metal futures in China using commodity futures contracts, and to Fung et al. (2003) which adopt the bivariate GARCH framework to analyze the information flow between commodity futures traded both in the US and China.

[3] See the document entitled *White Sugar Futures* (April 2009) on the Zhengzhou Commodity Exchange website http://www.czce.com.cn.

[4] By the end of 2013, there were 2.47 million investors trading in the futures market, 2.39 million of whom were individual investors (ChineseFuturesAssociation (2015, p.211)).

we are the first to offer solid and detailed evidence. Using five-minute returns data over long sample periods, we compute three popular liquidity measures that capture different aspects of liquidity, namely the effective spread of Roll (1984), the proportion of zero returns of Lesmond et al. (1999), and the Amihud (2002) illiquidity measure (Goyenko et al. (2009)). Our results show that contracts with three months to delivery are the most liquid as they exhibit the lowest effective spread, the lowest percentage of zero returns, and the smallest value for the Amihud (2002) illiquidity measure. This is different from the majority of futures markets and contracts for which the nearby contracts are usually the most liquid (see Baillie et al. (2007), Lee (2009), and the references therein). Crucially, this liquidity pattern results from the unique institutional environment in which trading takes place.

On the other hand, being an emerging market, the Chinese commodity futures market exhibits large proportion of zero returns (Bekaert et al. (2007)) and this is particularly evident in our five-minute return series. Even for the most liquid three-month to maturity contracts, the fraction of zero returns is as high as 36.27%, 23.90%, and 31.50% on average, respectively, for Aluminum, Copper, and Fuel Oil. In the existing literature, intraday data are widely adopted for volatility forecasting as they are shown to contain more information and provide more accurate and efficient forecasts (see, for example, Taylor and Xu (1997), Chortareas et al. (2011), Fuertes et al. (2015), and the references therein). However, the large proportion of zero returns in our data suggests that higher data sampling frequency does not necessarily translate into better forecasting performance due to information loss or noise in the data (Phillips and Yu (2009) and Bandi and Russel (2005)). Hence we choose to perform volatility forecasting by aggregating five-minute data into 15-, 30-, and 60-minute intraday returns and compute daily returns from daily prices so that we can observe and compare how well different models are at capturing the volatility dynamics given the data.

Equally important for the volatility forecast comparison is the choice of the true volatility proxy. While true volatility is a latent variable that cannot be observed in the market, an efficient and accurate representation of it is of great importance for the evaluation of volatility forecasts (see Andersen et al. (2010) for an excellent survey). In this paper, we undertake three different proxies for the true daily volatility. In addition to the widely adopted realized volatility measure of Andersen and Bollerslev (1998), we also consider the median-based measure of Andersen et al. (2012) and the range-based proxy advocated by Parkinson (1980), both of which are shown to be robust to zero returns, potential jumps in the underlying price dynamics, and other microstructure related effects (Alizadeh et al. (2002)).

In terms of volatility models, we begin with the conventional generalized autoregressive conditional heteroskedastic (GARCH) model of Bollerslev (1986, 1990). Our choice of models is also motivated by Baillie et al. (2007), which document strong long memory properties in commodity futures and argue that the fractionally integrated GARCH (FIGARCH) model captures this feature very well. At the same time, a natural alternative that works well at capturing the long memory property in realized volatility is the autoregressive fractionally integrated moving average (ARFIMA) model of Granger (1980) and Granger and Joyeux (1980). The two models differ in the manner in which information is extracted from intraday data: intraday returns are first aggregated to obtain daily realized volatility before the ARFIMA model is adopted to describe and forecast realized volatility at the daily level; whereas for the FIGARCH model, deseasonalized intraday data are directly fed into the model. So it is empirically interesting to compare the performance of the two models using our data.

Our empirical analysis reveals a host of interesting findings. First, in terms of the out-of-sample forecasting performance, the Diebold and Mariano (1995) and West (1996) test applied on a pairwise basis and the superior predicative ability test of Hansen (2005), which tests across alternative models simultaneously, suggest that the ARFIMA model consistently outperforms the GARCH-type models in the out-of-sample tests. It is the best performing model in 11 out of 15 commodity/volatility proxy combinations, and for the remaining four combinations the difference between the forecasting performance of the ARFIMA model and that of the best performing model is statistically insignificant at any conventional level. In other words, the ARFIMA model consistently produces the best forecasts or forecasts not inferior to the best in statistical terms.

It highlights the importance of incorporating the long memory dimension in volatility modeling in line with the literature. This finding also contributes to the discussion in the literature of whether the FIGARCH or the ARFIMA model is empirically better at capturing the long memory feature in the volatility dynamics (Chortareas et al. (2011)). Given that the intraday Chinese commodity futures data contain large proportion of zero returns which are directly fed in the FIGARCH model, it is not surprising that the ARFIMA model performs better.

Second, we show that within the GARCH family of models, the forecasting performance using the daily data is consistently as good as, if not better than, those using the intraday data. This finding suggests that the GARCH-type models may not be very efficient in utilizing the information contained in the intraday data of this particular market for volatility forecasting purpose due to high percentage of zero returns.

Finally, it is interesting to note that although Sugar contracts with January maturity and November maturity differ massively in terms of trading volume and show different levels of liquidity, the underlying volatility dynamics is nevertheless captured by the same model at the same data sampling frequency. For example, when the median- and range-based proxies are adopted, both futures contracts are best forecasted by the AFRIMA model using daily realized volatility obtained from the 60-minute returns. This further suggests that the ARFIMA model is a reliable and robust tool for forecasting volatility regardless of the underlying liquidity level with practical implications for traders and risk managers.

The rest of the paper is structured as follows. In Section 2, we briefly outline the alternative volatility models, the proxies for the true volatility dynamics, and the statistical metrics for the out-of-sample volatility forecasts evaluation. Section 3 describes the data and the model estimates. In Section 4, we discuss and analyze main empirical findings. Finally, Section 5 concludes. Details of the three liquidity measures are provided in the Appendix.

2 Models and Statistical Evaluation

2.1 Volatility Models

In this paper, we consider four popular volatility models at four different data sampling frequencies for volatility modeling and out-of-sample forecasting. In particular, we make use of the: (1) intraday GARCH, integrated GARCH (IGARCH), and FIGARCH models at the 15-, 30-, and 60-minute intervals; (2) daily GARCH, IGARCH, and FIGARCH models; and

(3) ARFIMA model applied to the daily realized volatility computed from the 15-, 30-, and 60-minute intervals. The model specifications are briefly outlined below.

GARCH Model

The GARCH model is the workhorse in the volatility estimation and forecasting literature (see, among others, Bollerslev (1986, 1990)). We use an ARMA(1,1) process in the conditional mean equation of the GARCH-type models. To allow for possible fat tails, we model the innovations in the GARCH process as independently and identically distributed Student's t-distribution while implementing the ARMA(1,1)-GARCH(1,1) model using both intraday and daily data. The model specification is given by

$$\begin{aligned} \tilde{r}_{t,n} &= \mu + \gamma \tilde{r}_{t,n-1} + \varepsilon_{t,n} + \theta \varepsilon_{t,n-1}, \quad \varepsilon_{t,n} | \Omega_{t,n-1} \sim D_v(0, h_{t,n}) \\ h_{t,n} &= \omega + \alpha \varepsilon^2_{t,n-1} + \beta h_{t,n-1}, \end{aligned} \tag{1}$$

where $\tilde{r}_{t,n}$ is the deseasonalized logarithmic return on day t for the nth time interval (see equations (10)-(12)), μ, γ, and θ are the parameters of the conditional mean equation, and ω, α, and β are the parameters of the conditional variance equation.[5] The error term $\varepsilon_{t,n}$, which is conditional on the information set $\Omega_{t,n-1}$, follows a Student's t-distribution (denoted by D_v) with zero mean, variance $h_{t,n}$, and v degrees of freedom. The GARCH model requires that $\alpha + \beta < 1$ for the volatility process to be stationary. For the IGARCH model, however, the corresponding requirement is $\alpha + \beta = 1$.

FIGARCH Model

The FIGARCH model extends the conditional variance equation of the standard GARCH model by adding fractional differences in order to allow for long memory property of the GARCH volatility process (Baillie et al. (1996) and Baillie and Morana (2009)). Following Baillie et al. (2000), we implement an ARMA(1,1)-FIGARCH(1,d,1) model given by

$$\tilde{r}_{t,n} = \mu + \gamma \tilde{r}_{t,n-1} + \varepsilon_{t,n} + \theta \varepsilon_{t,n-1}, \quad \varepsilon_{t,n} | \Omega_{t,n-1} \sim D_v(0, h_{t,n}) \tag{2}$$

[5] In case of daily data, r_t, h_t, ε_t, and Ω_{t-1} replace $\tilde{r}_{t,n}$, $\tilde{h}_{t,n}$, $\varepsilon_{t,n}$, and $\Omega_{t,n-1}$, respectively. Moreover, we do not deseasonalize daily returns used in the empirical analysis.

$$h_{t,n} \quad = \quad \omega + \beta h_{t,n-1} + [1 - \beta L_1 - (1 - \varphi L_1)(1 - L_1)^d]\varepsilon_{t,n}^2,$$

where ω, β, and φ are the parameters of the conditional variance equation, d is the order of fractional integration, L_1 is the lag operator on n, and D_v is the Student's t-distribution defined above.

ARFIMA Model

Granger (1980) and Granger and Joyeux (1980) introduce a flexible class of long memory processes based on realized volatilities not belonging to the ARCH family. It has been widely adopted in the literature when long memory properties are assumed in the data (see Martin and Wilkins (1999), Pong et al. (2003), and the references therein). The ARFIMA (p, d, q) model for a process y_t is defined as

$$\phi(L_2)(1 - L_2)^d(y_t - \mu) = \theta(L_2)\varepsilon_t, \tag{3}$$

where d is the order of fractional integration and L_2 is the lag operator on t. The AR and MA polynomial components are given as $\phi(L_2) = 1 + \phi_1 L_2 + \cdots + \phi_p L_2^p$ and $\theta(L_2) = 1 + \theta_1 L_2 + \cdots + \theta_q L_2^q$, respectively, and μ is the mean of y_t. In the empirical estimation of the ARFIMA (p, d, q) model, we follow Andersen et al. (2003) and replace y_t by the log of the daily realized volatility (denoted as $\log(\hat{\sigma}_t)$) obtained from the 15-, 30-, and 60-minute returns.

2.2 True Volatility Proxies

5-Minute Realized Volatility

The most popular proxy for the unobservable true volatility is the realized volatility measure proposed by Andersen and Bollerslev (1998). This is obtained by aggregating the intraday squared returns. We follow this approach and use a realized volatility series constructed from 5-minute log price series, which is the highest frequency in our data. The proxy is given by

$$\hat{\sigma}_{rv,t}^2 = \sum_{n=1}^{N} r_{t,n}^2, \tag{4}$$

where $\hat{\sigma}^2_{rv,t}$ is the realized variance for day t and $r^2_{t,n}$ is the squared 5-minute (log) return on day t for interval n $(n = 1, 2, \cdots, N)$.

Median-Based Volatility

The second proxy we exploit for true volatility is the median-based volatility measure introduced by Andersen et al. (2012). The measure is robust to jumps in the underlying return dynamics and to small ("zero") returns. The median-based true volatility proxy is defined as

$$\hat{\sigma}^2_{med,t} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{N}{N-2} \right) \times \sum_{n=2}^{N-1} \text{med}(|\Delta r_{n-1}|, |\Delta r_n|, |\Delta r_{n+1}|)^2, \tag{5}$$

where $\hat{\sigma}^2_{med,t}$ is the median-based variance for day t and $|\Delta r_n|$ is the absolute return over the nth interval on day t.

Range-Based Volatility

The third proxy for true volatility is the range-based measure proposed by Parkinson (1980). It has been further refined and adopted in Garman and Klass (1980), Yang and Zhang (2000), and Li and Hong (2011). Taking into account of daily high and low prices, this measure is able to deal with microstructure biases in the market. The proxy is defined as follows:

$$\hat{\sigma}^2_{rng,t} = \left(\frac{1}{4 \ln 2} (\ln H_t - \ln L_t) \right)^2, \tag{6}$$

where $\hat{\sigma}^2_{rng,t}$ is the range-based variance for day t , and H_t and L_t are the daily high and low prices, respectively.

2.3 Forecasting Accuracy

We use three different metrics to evaluate the out-of-sample forecasting accuracy of the volatility models, all of which are commonly adopted statistical measures in the literature (see, for example, Ahmed et al. (2016)).

Root Mean Squared Forecast Error

The root mean squared forecast error (RMSFE) compares the true volatility with the forecasted volatility from a given model and is computed as

$$\text{RMSFE} = \sqrt{\frac{1}{R}\sum_{t'=1}^{R}(\hat{h}_{t+1} - \hat{\sigma}^2_{t+1})^2}, \tag{7}$$

where R is the number of daily observations, $\hat{h}_{t+1}$ is the variance forecast, and $\hat{\sigma}^2_{t+1}$ is the chosen proxy for true variance in the out-of-sample period.

Diebold and Mariano (1995) and West (1996) Test

The second out-of-sample statistical metric of accuracy is the Diebold and Mariano (1995) and West (1996) MSFE t-statistic, which in our case tests whether a competing volatility model outperforms the benchmark volatility model by generating more accurate variance forecasts. We chose the benchmark model based on the lowest RMSFE. The test statistic is as follows:

$$\text{MSFE-}t = \frac{1}{\sqrt{R\hat{\Omega}}}\sum_{t=1}^{R}\Delta Loss_{t+1}, \tag{8}$$

where $\Delta Loss_{t+1}$ is the difference between the squared forecast error loss functions of the benchmark and competing volatility models and $\hat{\Omega}$ is the consistent estimate of the asymptotic variance of $R^{-0.5}\sum_{t=1}^{R}\Delta Loss_{t+1}$. The null hypothesis can be expressed as

$$\text{H}_0 : E[\Delta Loss_{t+1}] = 0. \tag{9}$$

Since the volatility models are non-nested, the alternative hypothesis in this case is two-sided. The test statistic in equation (12) follows an asymptotic standard normal distribution under the null hypothesis of equal predictive ability. We regress $\Delta Loss_{t'+1}$ on a constant and obtain the MSFE-t statistic for a zero coefficient based on the Andrews and Monahan (1992) estimator. A positive (negative) and statistically significant MSFE-t statistic suggests that the competing model outperforms (is outperformed by) the benchmark volatility model.

Superior Predictive Ability Test

To address the multiple-testing problem in the light of data mining, we conduct the superior predictive ability (henceforth SPA) test of Hansen (2005). Under the composite null hypothesis, there is no predictive ability across all competing volatility models. In other words, the null states that the benchmark model is not inferior to any of the alternative models. A rejection of the null hypothesis indicates that at least one competing model produces forecasts more accurate than the benchmark. Once again, we chose the benchmark model based on the lowest RMSFE and evaluate the out-of-sample forecasts based on the MSFE. For inference, we report stationary bootstrap p-values obtained using 10,000 replications.

3 Data and Estimation

The data come from the GTA Information Technology Company. We obtain contract ID, trading date, trading time, trading venue, contract expiry date, last recorded (Renminbi) price, high and low prices, and volume for 5-minute time series on four commodity futures contracts: Aluminum, Copper, Fuel Oil, and Sugar. The full sample period as well as the in-sample and out-of-sample periods for each commodity are provided in Table 1.[6,7] In Panel D, we find seasonality in trading volume for each contract over the full sample period. More precisely, we observe that in terms of average number of contracts traded for each delivery, there is not much variation across the 12 delivery months for Aluminum and Copper, and there is a slight variation for Fuel Oil. In other words, the number of contracts traded is relatively stable all-year round. However, with only six delivery months per year, Sugar shows a notable variation in the average number of contracts traded across the delivery months. In particular, contracts for January, May, and September exhibit huge trading volumes, while contracts for March, July, and November show the opposite. The trading volume for January delivery is the highest on average with more than 5.6 million contracts, whereas for November delivery the average trading volume is the lowest at 18418 contracts, about 0.32% of that for January delivery. This striking yet interesting variation naturally raises the question of how much the volatility dynamics for these two delivery months are different, if they are different at all. Hence, in the empirical exercises, we examine two futures contract series for Sugar, one for the very liquid January

[6] The starting and ending dates of the four commodities are constrained by data availability.

[7] Chortareas et al. (2011) and Liu et al. (2014) adopt similar sample period for the out-of-sample forecasting exercise with foreign exchange and commodity futures data, respectively.

delivery and the other for the very illiquid November delivery.

In Table 2, we report descriptive statistics of three measures adopted to describe liquidity of futures contracts at 5-minute interval, which is the highest sampling frequency in our data.[8] For Aluminum, the Roll spread measure for nearby contracts averages at 0.0006, zero returns account for 61% of all 5-minute returns on average in a trading day, and the scaled Amihud measure is 0.23. Comparing these figures to those for the three months to delivery contracts, we notice a marked improvement. In particular, the Roll spread drops to 0.0004, the percentage of zero returns decreases to 36%, and the scaled Amihud illiquidity measure drops to 0.03. The liquidity of the futures contract series subsequently worsens with longer time to delivery. For example, Aluminum contracts with three months to delivery are the most liquid and this liquidity decreases for contracts with longer or shorter time to maturity. The pattern is mirrored in the liquidity estimators for other commodities as well. Hence, in our volatility estimation and forecasting exercises for Aluminum, Copper, and Fuel Oil, we use futures contracts with three months to delivery, as they are the most liquid among all maturities, and volatility forecasts are least expected to be biased by the large proportion of zero returns.

While constructing the time series on returns with three months to maturity for Aluminum, Copper, and Fuel Oil, we choose prices of the third month prior to delivery month until the contract reaches the first day of two months prior to delivery month. We then switch to next contract, which is to be matured in three months to make continuous time series. Hence, for these three commodities, the contract time to maturity is always around three months. For Sugar futures, however, we are mostly interested in the effect that seasonality in trading volume has on volatility forecasting. Therefore, we take contracts from January to December for next January delivery and from November to October for next November delivery. This results in the contract time to maturity to change over time. The practice of switching contracts to the next delivery month is common in the literature (see, for example, Baillie et al. (2007)).

In our sample, all commodity futures are traded for four hours on a trading day starting at 9:00am and closing at 3:00pm with a two-hour break between 11:30 am and 1:30 pm. As a result, there are 48 5-minute returns on any business day. The (log) return $r_{t,n}$ on a trading

[8] A brief discussion of the three liquidity measures are contained in the Appendix.

day t for the nth interval is computed as

$$r_{t,n} = \ln P_{t,n} - \ln P_{t,n-1}, \tag{10}$$

where $P_{t,n}$ denote the commodity futures price on day t and the end of the nth interval. The 15-, 30-, 60-minute and daily returns are obtained by taking the logarithmic difference between prices that are 15, 30, and 60 minutes apart. The daily returns are computed as $r_t = \ln P_t - \ln P_{t-1}$.

In Table 3, we provide descriptive statistics of commodity futures contract returns at 5-, 15-, 30-, 60-minute and daily intervals. We notice that the average returns are very close to zero irrespective of contracts and data frequencies. Returns are left skewed with fat tails, although the degree of negative skewness and excess kurtosis tend to drop with decreasing sampling frequency. In addition, the percentage of zero returns drops considerably from the 5-minute to daily intervals. For example, it is 31.50% at the 5-minute interval, 17% at the 15-minute interval, while only 3.60% at the daily level for Fuel Oil. The trade-off between the improvement in data quality and the loss of information at lower frequencies could be crucial for the outcome of volatility measurement and forecasting exercises.

The volatility of intraday returns are known to display periodicity within a trading day, which could contaminate the estimation of conventional volatility models (Andersen and Bollerslev (1997)). Following Taylor and Xu (1997), we estimate a simple seasonality term $S_{t,n}$ by averaging the squared returns for each intraday period as follows:

$$\hat{S}_{t,n} = \frac{1}{T}\sum_{t=1}^{T} r_{t,n}^2, \tag{11}$$

where T is the number of trading days in the full sample period. The deseasonalized intraday returns are obtained as

$$\tilde{r}_{t,n} = \frac{r_{t,n}}{\hat{S}_{t,n}}. \tag{12}$$

We then make use of the deseasonlized returns to estimate the intraday GARCH family of models. In the out-of-sample forecasting, the intraday forecasts are based on the deseasonlized filtered returns and therefore transformed back to those from the original returns. This is implemented as follows:

$$\hat{h}_{t,n} = \hat{S}^2_{t,n} \times \tilde{h}_{t,n}, \tag{13}$$

where $\tilde{h}_{t,n}$ is the intraday variance forecast using the deseasonalized returns and $\hat{h}_{t,n}$ is the transformed variance forecast for the original returns. We produce one-step ahead daily volatility forecasts for daily models. But for intraday models, we produce 16-, 8-, and 4-step ahead forecasts for 15-, 30-, and 60-minute intervals and aggregate them to transform into daily forecasts. For the ARFIMA model, it is fitted directly to daily realized volatility aggregated from intraday returns. The out-of-sample forecasts are evaluated against the daily true volatility proxies described earlier. For all sampling frequencies, we use a rolling window forecasting scheme to obtain forecasts from all volatility models.

4 Empirical Analysis

4.1 In-Sample Results

We report the in-sample parameter estimates of the intraday GARCH, FIGARCH, and IGARCH models for five futures contracts at 15-, 30-, and 60-minute intervals in Table 4. For the ARMA(1,1)-GARCH(1,1) model specification in Panel A, most of the AR parameter estimates $\hat{\gamma}$ are statistically significant at conventional levels. Also, the MA parameter estimate $\hat{\theta}$ is significantly negative in most cases, capturing the first order negative autocorrelation in the returns. All the parameters in the conditional variance equations are highly significant at the 1% level except $\hat{\alpha}$ for 15-minute Copper contracts. The fact that $\hat{\alpha} + \hat{\beta} < 1$ reveals that the GARCH process is stationary, and, since $\hat{\alpha} + \hat{\beta}$ is close to 1, the volatility process is persistent. For the contract series with return innovations following a Student's t-distribution, the degrees of freedom parameter is between 2 and 4 and statistically significant at the 1% level. This indicates a fat tail in the return distributions.

In Panel B, when the volatility process is described by an ARMA(1,1)-FIGARCH(1,d,1) model, we notice that the parameter d, the order of fractional integration, is significantly different from zero at the 1% level for all futures contract series. This implies that the volatility process exhibits a long memory property and attests to the importance of adding this feature in the volatility dynamics of the commodity futures contract returns under scrutiny. It is also worth noting that, similar to the results in Panel A, the degrees of freedom parameter v is highly significant. Panel C shows the parameter estimates of the ARMA(1,1)-IGARCH(1,1) model specification and the results are qualitatively similar to those in Panel A.

Table 5 shows the in-sample parameter estimation for the daily GARCH, FIGARCH, and IGARCH models. These results are qualitatively similar to those in Table 4. We observe: (1) negative and significant first order autocorrelation in the conditional mean equation for each model and contract except for the daily IGARCH model using the Sugar contract with January delivery; (2) statistically significant $\hat{\beta}$ parameters; (3) highly significant fractional integration parameters $\hat{d}$; and (4) highly significant degrees of freedom parameters $\hat{v}$.

We present the in-sample parameter estimates of the ARFIMA model using the daily realized volatility obtained from the 15-, 30-, and 60-minute returns in Table 6. For Aluminum, Copper, and Fuel Oil, we set the MA term $q = 0$ as it is statistically insignificant at any conventional level. The first order autoregression term $\hat{p}$ is negative and highly significant and the fractional integration term $\hat{d}$ hovers around 0.4 for each of these three commodities. In cases of January and November contracts for Sugar, the first order autocorrelation $\hat{p}$ tends to be positive and quite often significant. The MA parameter $\hat{q}$ is close to -0.4 and significant at the 1% level. Similar to other commodities, the fractional integration parameter estimate for Sugar is in the vicinity of 0.45 and is highly significant.

Overall, the in-sample estimates of the GARCH, FIGARCH, IGARCH, and ARMIFA models reported in Tables 4 to 6 using intraday and daily data reveal that, for the four commodities, the return innovations are generally negatively autocorrelated with fat tails. Moreover, the underlying volatility processes are persistent with clear evidence of long memory properties.

4.2 Out-of-Sample Predictions

Table 7 reports RMSFEs for all volatility models, where forecasts errors are computed in comparison with three alternative true volatility proxies. In Panel A, we use the most widely exploited proxy in the literature, namely, the realized volatility measure constructed from the

5-minute returns. It is interesting to notice that for Aluminum and Copper futures contracts, the IGARCH and FIGARCH models produce the smallest RMSFEs, respectively, and both at the daily level. This preliminary evidence suggests that for this particular true volatility proxy, used in computing forecast errors, information contained in intraday prices does not help in generating more accurate volatility forecasts. For Fuel Oil, the 30-minute FIGARCH model produces the smallest RMSFE. It is also interesting to observe that although the January and November deliveries for Sugar contracts differ massively in terms of trading volume (see Table 1), the ARFIMA model utilizing the daily realized volatility obtained from the 15-minute returns provides the best forecasts for both futures contracts.

In Panel B, we consider median-based daily volatility as a proxy for true volatility. In this case, the ARFIMA model beats the rest of the competing models by producing the lowest RMSFE. More precisely, the ARFIMA model outperforms the other models for Copper, Fuel Oil, and Sugar (both January and November deliveries) when the daily realized volatility is obtained from the 60-minute returns. For Aluminum, it is the ARFIMA model using the daily realized volatility computed from the 30-minute returns. Finally, in Panel C, we make use of range-based volatility as true volatility proxy. Once again, the ARFIMA model is the best performing model for four out of five commodity futures contracts. In particular, the ARFIMA model applied to the daily realized volatility obtained from the 15-minute returns leads to the lowest RMSFE for Copper. But for Aluminum and January and November deliveries of Sugar contracts, it is the the 60-minute returns based daily realized volatility applied to the ARFIMA model. Fuel Oil is the only exception, for which the daily IGARCH model provides the most accurate out-of-sample variance forecasts.

Taken together, we notice three interesting and consistent patterns from the preliminary results in Table 7. First, the ARFIMA model, with its long memory dimension, dominates the other three volatility models in 11 out of 15 commodity/true volatility proxy combinations. Second, GARCH-type models using daily data outperform similar models using intraday data. Third, the ARFIMA model applied to the daily realized volatility obtained from the higher frequency returns (i.e., 15-minute returns) does not always beat the ARFIMA model using the daily realized volatility computed from the lower frequency returns. The latter two observations are novel for our chosen futures market because the literature seems to agree that intraday data enjoy informational advantage over daily data and that forecasting performance of the ARFIMA model improves with sampling frequency (Martens (2001) and Martens and Zein (2004)).

In Table 8, we provide pair-wise comparison following the well-known Diebold and Mariano (1995) and West (1996) test based on the Andrews and Monahan (1992) estimator. We choose the benchmark model in each case as the one with the lowest RMSFE in Table 7. The results suggest that the competing model forecasts are either as accurate statistically as the benchmark model, or, in most cases, significantly worse. It is interesting to notice that in Panel A, for Aluminum, the ARFIMA model utilizing the daily realized volatility from the 15-, 30-, ad 60-minute returns produces inferior forecasts but the difference from the benchmark is statistically insignificant. Put differently, the null hypothesis of equal MSFEs can not be rejected at any conventional level. In fact, for all model/true volatility proxy combinations, whenever the best performing model utilizes daily data, the ARFIMA model provides forecasts just as good statistically. These include the daily IGARCH model for Aluminum and the daily FIGARCH model for Copper in Panel A, and the daily IGARCH model for Fuel Oil in Panel C. For other model/true volatility proxy combinations, the competing models tend to produce statistically inferior forecasts, including both Sugar contracts in Panels A and C.

As a robustness check, we provide the Diebold and Mariano (1995) and West (1996) test results obtained by sequentially using each volatility model as the benchmark, based on their increasing RMSFEs, against the remaining alternative models in Tables A1 to A3. These additional results corroborate the conclusion in Table 8 that the benchmark, chosen as the one with the lowest RMSFE in Table 7, is indeed the one with the best volatility forecasting ability.

In Table 9, we perform the SPA test of Hansen (2005) to examine out-of-sample forecasting ability across all competing models and compute the stationary bootstrap p-values. The null hypothesis is that the benchmark model, the one with the lowest RMSFE, is not inferior to any of the competing models. The test results are resounding. The probability that the benchmark model is at least as good as the competing models in forecasting volatility in the out of sample is 1 or very close to it. Taken together, the results in Tables 8 and 9 clearly confirm and substantiate the observations in Table 7. In other words, when intraday data are directly used in the GARCH-type models, they are no better than daily data for volatility forecasting even after deseasonalization. Hence, if a model is to be recommended for volatility forecasting in the Chinese futures market, it would be the ARFIMA model, as it is consistently the best performing model or not inferior to the best performing one statistically.

Finally, we note that although Sugar contracts for January and November deliveries differ in terms of trading volume and liquidity, the underlying volatility dynamics is very similar. The in-sample parameter estimates are similar between these two series and both are best forecasted by the same model. When the 5-minute realized volatility is the proxy for true volatility, the ARFIMA model using the realized volatility computed from the 15-minute returns produces the most accurate forecast for both series, while the ARFIMA model applied to the realized volatility computed from the 60-minute interval outperforms competing models for the other two volatility proxies for both series. In other words, seasonality in trading volume and differences in liquidity do not affect volatility model selection.

5 Conclusion

In this paper, we undertake a comprehensive volatility forecasting exercise in a futures market with unique institutional regulations. In the Chinese commodity futures market, margin rate is time-dependent and investors face higher deposit as contracts move closer to maturity. In addition, although individuals account for the majority of investors, they are not allowed to trade nearby contracts. These two regulations result in a liquidity pattern whereby contracts with three months to delivery are the most liquid and we demonstrate this by computing three popular liquidity measures with 5-minute intraday data for Aluminum, Copper, Fuel Oil, and Sugar. In addition, even these most liquid contract series contain large percentage of zero returns at the 5-minute interval.

We explicitly take these features into account when forecasting volatility and utilize more distant three months to maturity contracts at the daily and three different intraday sampling frequencies. We demonstrate that the long memory dimension is present in our data in the in-sample volatility modeling. When it comes to out-of-sample forecasting, we show that the

ARFIMA model, which aggregates intraday returns to daily level in generating daily forecasts, is the best-performing model, or equivalent to the best-performing model in statistical terms. The FIGARCH model, which also incorporates the long memory feature in the volatility dynamics, is less efficient in generating forecasts probably due to the fact that large proportions of intraday returns are zero and the deseasonalized intraday returns are direct fed into the model.

Furthermore, we show that within the GARCH-family of models, the forecasting performance using the daily data is consistently as good as, if not better than, those using the intraday data, which also attests to the trade-off between information and noise in the intraday data with many zero returns. Finally, it is interesting to note that even though January and November contract series for Sugar differ massively in terms of trading volume, their underlying volatility dynamics are well captured and forecasted by the ARFIMA model at the same data sampling frequency.

Acknowledgement

We thank comments and suggestions by participants at the 2015 Asian Financial Association annual conference and the Workshop on Chinese Commodity Futures Market. Thanks are also due to seminar audience at Renmin University of China. Jiang and Liu gratefully acknowledge financial support from the Humanities and Social Sciences Research Fund for Young Scientists by the Ministry of Education of China (Grant No. 12YJC790079).

Appendix: Liquidity Measures

We use three liquidity estimators widely adopted in the literature to describe the liquidity of the Chinese commodity futures contracts. They are the effective spread of Roll (1984), the proportion of zero returns as in Lesmond et al. (1999), and the Amihud (2002) illiquidity estimator. These measures are shown to perform quite well in capturing the different aspects of the asset liquidity (Goyenko et al. (2009)).

Roll Spread

In the seminal paper of Roll (1984), a simple serial covariance spread estimation model is developed to capture asset liquidity. The effective spread is derived from the serial covariance properties of transaction price changes. The model has led to a burgeoning research area in the market microstructure literature with many modifications and extensions (see George et al. (1991), Chang and Chang (1993), and the references therein).

To illustrate, let E and P_t denote the effective spread and the closing price on day t, respectively, and Δ is the change operator. Roll (1984) shows that the serial covariance between changes in prices is

$$E = 2\sqrt{-\mathrm{Cov}(\Delta P_t, \Delta P_{t-1})}. \tag{A1}$$

In this paper, we follow Goyenko et al. (2009) and adopt a modified version of the Roll (1984) spread so that we can always obtain a numerical value for this liquidity measure. Denoting the price change over the nth time interval as ΔP_n, the effective spread can be expressed as follows:

$$\mathrm{Roll} = \begin{cases} 2\sqrt{-\mathrm{Cov}(\Delta P_n, \Delta P_{n-1})} & \text{if } \mathrm{Cov}(\Delta P_n, \Delta P_{n-1}) < 0 \\ 0 & \text{otherwise} \end{cases}. \tag{A2}$$

Hence, the lower the effective spread, the higher the liquidity of the asset.

Proportion of Zero Returns

The second liquidity measure we exploit is proposed in Lesmond et al. (1999) and proves especially useful and effective in studying liquidity of emerging markets (see, among others, Bekaert et al. (2007) and Lesmond (2005)). This measure is based on the transaction cost, that is, if the value of an information signal is insufficient to outweigh the cost associated with trading, market participants will choose not to trade, resulting in a zero return. The measure is easy to implement since it only requires a time series on transaction data. In this paper, the proportion of zero returns in a trading day is defined as follows:

$$\mathrm{Zeros} = (\#\text{ of intraday time intervals with zero returns})/N, \tag{A3}$$

where N is the total number of time intervals in a trading day ($n = 1, 2, \cdots, N$). Intuitively, the lower is the proportion of zero returns, the better is the liquidity of the asset.

Amihud Illiquidity Measure

The illiuqidity measure of Amihud (2002) is another popular estimator in the literature (see, among others, Baker and Stein (2004) and Amihud et al. (2012)). It is a price impact measure that captures the price response associated with one unit currency of trading volume. Hence, the lower is the illiquidity measure, the better is the asset liquidity. More precisely, it is defined as the ratio given by

$$\text{Amihud} = \text{Average}\left(\frac{|r_n|}{\text{Volume}_n}\right), \tag{A4}$$

where r_n is the asset return in log over the nth time interval and Volume_n is the US dollar (in our case, Renminbi) trading volume over the same interval.

References

Ahmed, S., Liu, X., Valente, G., 2016. Can currency-based risk factors help forecast exchange rates? International Journal of Forecasting 32, 75–97.

Alizadeh, S., Brandt, M. W., Diebold, F. X., 2002. Range-based estimation of stochastic volatility models. Journal of Finance 57, 1047–1091.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Amihud, Y., Mendelson, H., Pedersen, L. H., 2012. Market Liquidity: Asset Pricing, Risk, and Crises. Cambridge University Press.

Andersen, T., Bollerslev, T., Diebold, F., 2010. Parametric and nonparametric volatility measurement. In: Ait-Sahalia, Y., Hansen, L. (Eds.), Handbook of Financial Econometrics: Tools and Techniques. North-Holland/Elsevier, Amsterdam; Boston, pp. 67–137.

Andersen, T. G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T. G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Andersen, T. G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169, 75–93.

Andrews, D. W. K., Monahan, C. J., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60, 953–966.

Baillie, R. T., Bollerslev, T., Mikkelsen, H.-O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3–30.

Baillie, R. T., Cencen, A. A., Han, Y. W., 2000. High frequency Deutsche Mark-US Dollar returns: FIGARCH representations and non-linearities. Multinational Finance Journal 4, 247–267.

Baillie, R. T., Han, Y.-W., Myers, R. J., Song, J., 2007. Long memory models for daily and high frequency commodity futures returns. Journal of Futures Markets 27, 643–668.

Baillie, R. T., Morana, C., 2009. Modeling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach. Journal of Economic Dynamics & Control 33, 1577–1592.

Baker, M., Stein, J. C., 2004. Market liquidity as a sentiment indicator. Journal of Financial Markets 7, 271–299.

Bandi, F. M., Russel, J. R., 2005. Separating microstructure noise from volatility. Journal of Financial Economics 79, 655–692.

Bekaert, G., Harvey, C. R., Lundblad, C., 2007. Liquidity and expected returns: Lessons from emerging markets. Review of Financial Studies 20, 1783–1831.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., 1990. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. Review of Economics and Statistics 72, 498–505.

Chang, C. W., Chang, S. K., 1993. An implicit measure of the effective bid-ask spread: A note. Journal of Financial Research 16, 71–75.

ChineseFuturesAssociation, 2015. Convoy in the Futures Markets: Protecting the rights of futures investors (in Chinese). China Finance Press.

Chortareas, G., Jiang, Y., Nankervis, J. C., 2011. Forecasting exchange rate volatility using high-frequency data: Is the euro different? International Journal of Forecasting 27, 1089–1107.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Fuertes, A. M., Kalotychou, E., Todorovic, N., 2015. Daily volume, intraday and overnight returns for volatility prediction: Profitability or accuracy? Review of Quantitative Finance and Accounting 45, 251–278.

Fung, H.-G., Leung, W. K., Xu, X., 2003. Information flows between the US and China commodity futures trading. Review of Quantitative Finance and Accounting 21, 267–285.

Garman, M. B., Klass, M. J., 1980. On the estimation of security price volatilities from historical data. Journal of Business 53, 67–78.

George, T. J., Kaul, G., Nimalendran, M., 1991. Estimation of the bid-ask spread and its components: A new approach. Review of Financial Studies 4, 623–656.

Goyenko, R. Y., Holden, C. W., Trzcinka, C. A., 2009. Do liquidity measures measure liquidity. Journal of Financial Economics 92, 153–181.

Granger, C. W., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227–238.

Granger, C. W., Joyeux, R., 1980. An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15–39.

Hansen, P. R., 2005. A test of superior predictive ability. Journal of Business & Economic Statistics 23, 365–380.

Lee, H.-T., 2009. Optimal futures hedging under jump switching dynamics. Journal of Empirical Finance 16, 446–456.

Lesmond, D. A., 2005. The costs of equity trading in emerging markets. Journal of Financial Economics 77, 411–452.

Lesmond, D. A., Ogden, J. P., Trzcinka, C. A., 1999. A new estimate of transaction costs. Review of Financial Studies 12, 1113–1141.

Li, H., Hong, Y., 2011. Financial volatility forecasting with range-based autoregressive volatility model. Finance Research Letters 8, 69–76.

Liu, Q., Chng, M., Xu, D., 2014. Hedging industrial metals with stochastic volatility models. Journal of Futures Markets 34, 704–730.

Martens, M., 2001. Forecasting daily exchange rate volatility using intraday returns. Journal of International Money and Finance 20, 1–23.

Martens, M., Zein, J., 2004. Predicting financial volatility: High-frequency time-series forecasts vis-a-vis implied volatility. Journal of Futures Markets 24, 1005–1028.

Martin, V. L., Wilkins, N. P., 1999. Indirect estimation of ARFIMA and VARFIMA models. Journal of Econometrics 93, 149–175.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. Journal of Business 53, 61–65.

Peck, A. E., 2008. The development of futures markets in China: Evidence of some unique trading characteristics. In: Goss, B. (Ed.), Debt, Risk, and Liquidity in Futures Markets. Routledge, Oxon; New York, pp. 46–74.

Phillips, P. C. B., Yu, J., 2009. Information loss in volatility measurement with flat price trading, Unpublished manuscript, Yale University and Singapore Management University.

Pong, S., Shackleton, M., Taylor, S. J., Xu, X., 2003. Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. Journal of Banking and Finance 28, 2541–2563.

Poon, S.-H., Granger, C., 2003. Forecasting financial market volatility: A review. Journal of Economic Literature 41, 478–539.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance 39, 1127–1139.

Taylor, S. J., Xu, X., 1997. The incremental volatility information in one million foreign exchange quotations. Journal of Empirical Finance 4, 317–340.

West, K. D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.

Yang, D., Zhang, Q., 2000. Drift-independent volatility estimation based on high, low, open, and close prices. Journal of Business 73, 477–492.

Table 1. Sample Periods and Trading Volumes for Commodity Futures Contracts

The table presents the full sample periods, the in-sample periods, and the out-of-sample periods, respectively, in Panels A to C for Aluminum, Copper, Fuel Oil, and Sugar. Panel C reports the number of trading days for the out-of-sample forecasts. Panel D reports the average number of contracts traded for each delivery month over the full sample period for each commodity.

	Aluminum	Copper	Fuel Oil	Sugar
		Panel A: Full Sample Period		
From	1 Aug 2003	1 Aug 2003	8 Oct 2004	6 Jan 2006
To	19 Dec 2013	19 Dec 2013	30 Sep 2011	14 Jul 2014
		Panel B: In-Sample Period		
From	1 Aug 2003	1 Aug 2003	8 Oct 2004	6 Jan 2006
To	17 Sep 2012	17 Sep 2012	8 Dec 2010	17 Apr 2013
		Panel C: Out-of-Sample Period		
From	18 Sep 2012	18 Sep 2012	9 Dec 2010	18 Apr 2013
To	19 Dec 2013	19 Dec 2013	30 Sep 2011	14 Jul 2014
No. of days	300	300	200	300
		Panel D: Trading Volume		
Jan	144825	546380	238806	5686023
Feb	109620	452251	513169	N/A
Mar	154988	420790	396213	296452
Apr	114904	297649	24687	N/A
May	138448	357730	341555	4460179
Jun	115161	364373	192583	N/A
Jul	117022	392841	197663	300749
Aug	104490	520152	130340	N/A
Sep	98125	611807	162952	4343036
Oct	132359	635110	117432	N/A
Nov	156022	592573	175998	18418
Dec	125845	557593	176067	N/A

Table 2. Liquidity Measures of Commodity Futures with Different Time to Delivery

The table reports descriptive statistics of liquidity for Aluminum, Copper, Fuel Oil, and Sugar contracts at 5-minute interval using three liquidity measures. Roll refers to the effective spread of Roll (1984) ($\times 10^3$); Zeros are the proportion of 5-minute zero returns during a trading day; and Amihud is the illiquidity measure of Amihud (2002) ($\times 10^8$). The futures contracts are grouped according to their time to delivery. The full sample period for each commodity futures contract is reported in Table 1.

	Measure	Mean	Median	Stdev	Max	Mean	Median	Stdev	Max	Mean	Median	Stdev	Max	Mean	Median	Stdev	Max	Mean	Median	Stdev	Max
		Aluminum				Copper				Fuel Oil				Sugar (Jan)				Sugar (Nov)			
Nearby Month	Roll	0.5777	0.2921	1.0067	13.124	0.7238	0.4502	1.1667	19.511					1.5233	0.3903	2.7390	13.045	0.9875	0.3056	2.0844	12.830
	Zeros	0.6127	0.6042	0.1513	1	0.4231	0.3958	0.1761	1		N/A			0.8191	0.8750	0.1477	1	0.8713	0.8750	0.0809	1
	Amihud	0.2257	0.1215	0.3854	5.8025	0.1599	0.0489	0.7954	20.510					35.860	9.1569	64.969	307.82	32.890	10.712	60.801	311.77
One Month	Roll	0.4814	0.3053	0.6842	10.989	0.6156	0.4723	0.7222	7.3547	0.9800	0.3035	2.2287	28.729	0.6433	0.5001	0.8231	8.05461	1.4276	0.8149	2.1002	11.665
	Zeros	0.5014	0.4898	0.1797	1	0.3126	0.2500	0.1805	1	0.8883	0.9375	0.1184	1	0.3819	0.3265	0.2202	0.9583	0.5385	0.5000	0.2366	1
	Amihud	0.2599	0.1329	0.4998	9.774	0.1557	0.0266	1.2068	31.872	8.5623	2.7080	21.744	413.06	3.5725	1.0095	6.1811	35.067	12.718	3.4806	24.377	147.31
Two Months	Roll	0.4533	0.3109	0.5902	6.2196	0.5465	0.4319	0.6573	7.7768	0.7454	0.4811	1.0630	11.700	0.6643	0.5137	0.6460	2.9291	0.9704	0.7481	1.0066	5.4797
	Zeros	0.4302	0.3958	0.1809	1	0.2713	0.2083	0.1883	1	0.3848	0.2916	0.2551	1	0.2704	0.2083	0.1832	1	0.4419	0.3438	0.2490	0.9375
	Amihud	0.0650	0.0290	0.1102	1.2475	0.0856	0.0042	0.7167	24.478	6.3031	0.2223	28.219	499.11	1.8920	0.1000	4.3022	22.3158	11.7065	1.4452	21.8190	139.26
Three Months	Roll	0.4413	0.3254	0.5783	7.0231	0.5457	0.4326	0.6747	12.373	0.5036	0.3525	0.6446	6.9974	0.6666	0.5228	0.7613	4.2882	0.9113	0.7279	0.8884	4.6090
	Zeros	0.3627	0.3750	0.1713	1	0.2390	0.1875	0.1837	1	0.3150	0.2500	0.2509	1	0.2138	0.1667	0.1497	1	0.3531	0.2500	0.2472	1
	Amihud	0.0294	0.0088	0.0495	0.5031	0.0514	0.0015	0.2359	3.2224	2.4752	0.0218	16.490	528.57	0.9727	0.0187	2.8976	19.049	7.9827	0.6745	17.435	80.557
Four Months	Roll	0.4728	0.3079	1.5747	73.105	0.6013	0.4806	0.6899	6.4027	0.6401	0.4381	0.9000	13.095	0.6443	0.5608	0.6559	2.9074	0.6675	0.5026	0.7754	4.5419
	Zeros	0.4260	0.3958	0.1777	1	0.3015	0.2292	0.1970	1	0.4624	0.3542	0.2786	1	0.2113	0.1667	0.1223	0.7708	0.4568	0.3646	0.2606	0.9375
	Amihud	0.1658	0.0679	0.5201	16.169	0.1174	0.0138	0.6839	31.676	6.2857	1.5769	21.792	655.46	0.9772	0.0055	2.6047	11.879	6.5388	1.0196	11.693	60.866
Five Months	Roll	0.4855	0.2988	0.6870	11.378	0.6786	0.5084	0.8849	17.583	0.7476	0.3781	1.1416	10.839	0.6507	0.5691	0.6631	3.9315	0.6516	0.4816	0.8682	5.7352
	Zeros	0.5194	0.5000	0.2116	1	0.3996	0.3125	0.2392	1	0.6479	0.6875	0.2603	1	0.2027	0.1837	0.1094	1	0.4798	0.4167	0.2218	0.9792
	Amihud	0.7390	0.3192	1.7928	24.401	0.3748	0.1378	1.3640	33.787	18.728	8.5319	40.796	631.32	0.4032	0.0016	1.3262	9.2732	5.7274	2.0260	8.6239	52.610

Table 3. Descriptive Statistics of Commodity Futures Returns

The table reports descriptive statistics of commodity futures returns at 5-, 15-, 30-, 60-minute and daily intervals. For Aluminum, Copper, and Fuel Oil, we choose prices of the third month prior to delivery month until the contract reaches the first day of two months prior to delivery month. Then we switch to next contract, which is to be matured in three months to make continuous time series. For Sugar, we choose contracts with the most (January) and the least (November) liquid delivery months. At each January (November) contract delivery month, the data switch to the January (November) contract maturing in the following year. The full sample period for each commodity futures contract is reported in Table 1.

Commodity	Interval	Mean	Stdev	Skew	Kurt	Min	Max	Count	Zero Return
Aluminum	5-min	−2.9E-06	0.002	−2.960	180.469	−0.056	0.046	119357	36.27%
	15-min	−8.7E-06	0.003	−1.695	64.442	−0.055	0.046	39982	22.78%
	30-min	−1.6E-05	0.004	−1.232	34.453	−0.055	0.046	20230	15.38%
	60-min	−3.4E-05	0.005	−0.943	18.404	−0.058	0.046	10334	10.42%
	Daily	−1.2E-04	0.010	−0.602	4.573	−0.060	0.041	2521	0.04%
Copper	5-min	1.28E-05	0.003	−1.291	126.346	−0.062	0.064	120606	23.90%
	15-min	3.84E-05	0.004	−0.608	42.690	−0.062	0.063	40478	12.50%
	30-min	7.68E-05	0.006	−0.388	21.504	−0.062	0.066	20446	8.16%
	60-min	1.49E-04	0.008	−0.296	10.161	−0.062	0.068	10438	3.30%
	Daily	5.92E-04	0.016	−0.226	1.364	−0.062	0.057	2522	1.98%
Fuel Oil	5-min	1.05E-05	0.002	−2.071	121.288	−0.061	0.056	74160	31.50%
	15-min	3.16E-05	0.004	−1.196	43.084	−0.061	0.055	24720	17.00%
	30-min	6.13E-05	0.005	−0.848	21.564	−0.061	0.058	12360	11.70%
	60-min	1.24E-04	0.008	−0.676	10.372	−0.061	0.059	6172	7.00%
	Daily	5.36E-04	0.015	−0.268	2.249	−0.059	0.058	1544	3.60%
Sugar (Jan)	5-min	−1.40E-06	0.002	−1.570	148.205	−0.078	0.058	98661	21.84%
	15-min	−3.95E-06	0.003	−0.961	57.962	−0.078	0.058	33253	11.81%
	30-min	−7.84E-06	0.005	−0.782	31.775	−0.078	0.058	16901	7.00%
	60-min	−9.34E-06	0.006	0.012	27.693	−0.079	0.116	8725	5.20%
	Daily	1.64E-05	0.013	−0.050	2.478	−0.078	0.058	2046	1.00%
Sugar (Nov)	5-min	−4.00E-07	0.002	−0.448	115.778	−0.078	0.053	98556	55.60%
	15-min	−1.30E-06	0.003	−0.413	44.935	−0.078	0.055	33212	34.92%
	30-min	−2.90E-06	0.005	−0.161	24.756	−0.078	0.055	16877	24.27%
	60-min	−8.31E-06	0.006	−0.264	13.689	−0.078	0.053	8707	16.97%
	Daily	−4.63E-05	0.012	−0.045	2.935	−0.075	0.058	2037	1.70%

Table 4. In-Sample Parameter Estimation of the Intraday GARCH, FIGARCH, and IGARCH Models

The table reports the in-sample parameter estimates of the intraday GARCH, FIGARCH, and IGARCH models. In all panels, estimates are obtained using 15-, 30-, and 60-minute deseasonalized intraday returns. The models are estimated using quasi-maximum likelihood with Student's t-distributed innovations with v degrees of freedom. Only for Fuel Oil, the GARCH model at 15-minute interval and for Sugar (November), the GARCH, FIGARCH, and IGARCH models at 15-, 30-, and 60-minute intervals are estimated assuming a normal distribution. Numbers in parentheses are t-statistics, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The in-sample period for each commodity futures contract is reported in Table 1.

	Aluminum			Copper			Fuel Oil			Sugar (Jan)			Sugar (Nov)		
	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min
	Panel A: ARMA(1,1)-GARCH(1,1)														
$\hat{\gamma}$	0.36***	0.34***	0.04	0.06	0.23**	−0.14**	−0.79***	0.34***	0.12	0.18***	0.07	−0.18	−0.08	0.24***	−0.17**
	(7.15)	(7.92)	(0.67)	(0.86)	(2.41)	(−2.19)	(−7.00)	(3.42)	(0.18)	(3.47)	(0.28)	(−1.37)	(−0.54)	(2.61)	(−2.08)
$\hat{\theta}$	−0.46***	−0.43***	−0.13***	−0.12*	−0.25***	0.11*	0.78***	−0.39***	−0.15	−0.26***	−0.10	0.14	0.02	−0.28***	0.10
	(−9.54)	(−10.38)	(−2.45)	(−1.82)	(−2.71)	(1.78)	(6.62)	(−3.92)	(−0.22)	(−5.15)	(−0.43)	(1.07)	(0.11)	(−3.08)***	(1.23)
$\hat{\alpha}$	0.06***	0.11***	0.08***	0.10	0.04***	0.05***	0.05***	0.05***	0.06***	0.05***	0.04***	0.04***	0.02***	0.03***	0.05***
	(7.07)	(8.22)	(6.07)	(0.88)	(8.60)	(8.95)	(5.84)	(6.29)	(7.61)	(6.94)	(5.27)	(4.54)	(5.95)	(5.12)	(5.57)
$\hat{\beta}$	0.93***	0.89***	0.91***	0.90***	0.95***	0.94***	0.94***	0.95***	0.94***	0.95***	0.95***	0.95***	0.97***	0.96***	0.94***
	(98.53)	(124.80)	(67.84)	(9.02)	(213.40)	(168.10)	(101.50)	(123.60)	(145.90)	(151.10)	(126.00)	(108.10)	(268.1)	(137.40)	(84.37)
$\hat{v}$	2.67***	2.66***	2.67***	2.78***	2.89***	3.24***		3.48***	3.86***	2.78***	2.88***	3.11***			
	(54.17)	(40.63)	(29.02)	(20.82)	(38.16)	(25.53)		(26.95)	(17.72)	(50.69)	(36.09)	(23.91)			
	Panel B: ARMA(1,1)-FIGARCH(1,d,1)														
$\hat{\gamma}$	0.35***	0.34***	0.04	0.09	0.21**	−0.16**	0.36***	0.36***	−0.02	0.17***	0.07	−0.18	0.03	0.26**	−0.17*
	(5.64)	(7.80)	(0.74)	(1.37)	(2.07)	(−2.36)	(4.41)	(3.44)	(−0.14)	(3.59)	(0.30)	(−1.33)	(0.18)	(2.56)	(−1.94)
$\hat{\theta}$	−0.44***	−0.43***	−0.13**	−0.16**	−0.24***	0.13**	−0.43***	−0.40***	−0.06	−0.26***	−0.10	0.15	−0.09	−0.30***	0.10
	(−7.36)	(−10.22)	(−2.46)	(−2.27)	(−2.30)	(1.97)	(−5.54)	(−3.87)	(−0.43)	(−5.46)	(−0.46)	(1.05)	(−0.55)	(−3.00)	(1.17)
$\hat{\beta}$	0.70***	0.77***	0.67***	0.82***	0.82***	0.83***	0.70***	0.82***	0.78***	0.81***	0.81***	0.71***	0.76***	0.76***	0.77***
	(18.90)	(31.05)	(16.17)	(21.19)	(42.37)	(9.82)	(18.96)	(31.17)	(18.10)	(40.67)	(37.99)	(15.51)	(13.93)	(16.40)	(9.45)
$\hat{\varphi}$	0.29***	0.46***	0.29***	0.72***	0.53***	0.24**	0.42***	0.45***	0.36***	0.59***	0.45***	0.29***	0.56***	0.50***	0.38***
	(12.03)	(14.40)	(8.13)	(13.17)	(15.78)	(2.44)	(11.27)	(11.88)	(8.60)	(20.10)	(13.16)	(6.72)	(8.07)	(8.31)	(5.71)
$\hat{d}$	0.38***	0.52***	0.57***	0.36***	0.48***	0.67***	0.43***	0.56***	0.54***	0.42***	0.51***	0.47***	0.33***	0.42***	0.54***
	(24.21)	(17.03)	(13.37)	(20.32)	(14.73)	(3.65)	(18.49)	(10.49)	(9.31)	(18.84)	(12.97)	(8.31)	(13.96)	(9.53)	(4.13)

(Continued)

Table 4 – *Continued*

	Aluminum			Copper			Fuel Oil			Sugar (Jan)			Sugar (Nov)		
	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min	15-min	30-min	60-min
$\hat{v}$	2.67***	3.22***	3.30***	3.23***	3.26***	3.52***	3.75***	3.83***	3.76***	3.15***	3.16***	3.32***			
	(54.17)	(62.38)	(43.79)	(67.77)	(48.24)	(26.40)	(46.69)	(31.62)	(24.02)	(66.64	(47.09)	(29.74)			
	Panel C: ARMA(1,1)-IGARCH(1,1)														
$\hat{\gamma}$	0.35***	0.35***	0.04	0.05	0.23**	−0.14**	0.36***	0.35***	−0.01	0.17***	0.07	0.69***	−0.04	0.24***	−0.17**
	(6.64)	(7.78)	(0.73)	(0.74)	(2.37)	(−2.19)	(4.35)	(3.37)	(−0.06)	(3.33)	(0.28)	(3.92)	(−0.30)	(2.64)	(−2.14)
$\hat{\theta}$	−0.45***	−0.43***	−0.13**	−0.11*	−0.25***	0.11*	−0.43***	−0.39***	−0.02	−0.25***	−0.10	−0.70***	−0.02	−0.28***	0.11
	(−8.87)	(−10.12)	(−2.42)	(−1.68)	(−2.65)	(1.80)	(−5.44)	(−3.86)	(−0.14)	(−4.95)	(−0.42)	(−4.15)	(−0.19)	(−3.11)	(1.27)
$\hat{\alpha}$	0.07***	0.07***	0.09***	0.09	0.04***	0.05***	0.06***	0.05***	0.06***	0.05***	0.04***	0.04***	0.02***	0.04***	0.06***
	(8.36)	(10.27)	(7.35)	(1.26)	(9.05)	(9.85)	(7.58)	(6.52)	(8.27)	(7.36)	(5.51)	(4.62)	(2.99)	(5.04)	(5.54)
$\hat{v}$	3.14***	3.10***	3.22***	2.94***	3.09***	3.45***	3.52***	3.62***	3.97***	3.00***	3.04***	3.18***			
	(84.89)	(62.87)	(42.90)	(31.15)	(51.86)	(9.58)	(48.87)	(33.34)	(20.67)	(70.80)	(49.33)	(31.95)			

Table 5. In-Sample Parameter Estimation of the Daily GARCH, FIGARCH, and IGARCH Models

The table reports the in-sample parameter estimates of the daily GARCH, FIGARCH, and IGARCH models. The models are estimated using quasi-maximum likelihood with Student's t-distributed innovations with v degrees of freedom. Numbers in parentheses are t-statistics, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The in-sample period for each commodity futures contract is reported in Table 1.

Model	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\varphi}$	$\hat{d}$	$\hat{v}$
Panel A: Aluminium							
GARCH	0.78 (5.65)***	−0.80 (−6.03)***	0.23 (6.59)***	0.76 (29.61)***			3.81 (11.28)**
FIGARCH	0.80 (6.55)***	−0.83 (−7.04)***		0.67 (10.18)***	0.21 (3.01)***	0.70 (8.24)***	4.52 (13.49)***
IGARCH	−0.50 (−2.19)**	0.51 (2.29)**	0.19 (8.04)***				4.50 (13.31)***
Panel B: Copper							
GARCH	0.48 (3.15)***	−0.42 (−2.71)***	0.10 (7.16)**	0.88 (53.43)***			9.14 (4.91)***
FIGARCH	0.46 (2.91)***	−0.40 (−2.49)***		0.59 (3.33)***	0.12 (1.56)	0.53 (3.57)***	8.76 (4.73)***
IGARCH	0.48 (3.13)***	−0.42 (−2.72)***	0.11 (6.82)***				7.74 (5.21)***
Panel C: Fuel Oil							
GARCH	0.22 (1.10)	−0.33 (−1.78)*	0.08 (5.02)***	0.91 (58.01)***			5.91 (6.23)***
FIGARCH	0.21 (1.14)	−0.33 (−1.86)*		0.84 (14.15)***	−0.01 (−0.09)	0.86 (7.12)***	5.98 (6.65)***
IGARCH	0.22 (1.11)	−0.33 (−1.79)*	0.09 (5.38)***				5.78 (6.71)***
Panel D: Sugar (Jan)							
GARCH	−0.92 (−13.44)***	0.93 (13.92)***	0.12 (6.13)***	0.88 (49.41)***			5.76 (7.18)***
FIGARCH	−0.93 (−18.82)***	0.93 (19.33)***		0.83 (7.29)***	0.10 (0.66)	0.87 (3.34)***	5.77 (6.70)***
IGARCH	−0.63 (−0.37)	0.64 (0.37)	0.12 (6.68)***				5.48 (6.30)***
Panel E: Sugar (Nov)							
GARCH	−0.77 (−3.37)***	0.78 (3.43)***	0.11 (5.58)***	0.89 (48.90)***			4.67 (8.85)***
FIGARCH	−0.82 (−2.41)**	0.83 (2.49)**		0.86 (9.66)***	0.14 (1.00)	0.89 (4.06)***	4.81 (7.63)***
IGARCH	−0.77 (−3.38)***	0.78 (3.44)***	0.11 (6.13)***				4.65 (9.03)***

Table 6. In-Sample Parameter Estimation of the ARFIMA(p, d, q) Model

The table reports the in-sample parameter estimates of the ARFIMA(p, d, q) model using the daily realized volatility computed from the 15-, 30-, and 60-minute returns. Numbers in parentheses are t-statistics, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The in-sample period for each commodity futures contract is reported in Table 1.

Commodity	Return Interval	AR(1)	MA(1)	d
Aluminum	15-min	−0.14 (−5.15)***		0.47 (24.40)***
	30-min	−0.16 (−5.74)***		0.47 (21.60)***
	60-min	−0.12 (−3.97)***		0.38 (17.80)***
Copper	15-min	−0.20 (−7.15)***		0.41 (19.40)***
	30-min	−0.22 (−8.15)***		0.40 (19.20)***
	60-min	−0.22 (−8.11)***		0.37 (17.60)***
Fuel Oil	15-min	−0.22 (−6.55)***		0.38 (15.60)***
	30-min	−0.23 (−6.77)***		0.37 (14.80)***
	60-min	−0.20 (−5.73)***		0.33 (13.50)***
Sugar (Jan)	15-min	0.20 (1.95)**	−0.40 (−3.87)***	0.48 (21.00)***
	30-min	0.20 (2.30)**	−0.45 (−4.94)***	0.48 (16.30)***
	60-min	0.12 (1.25)	−0.41 (−3.49)***	0.45 (10.50)***
Sugar (Nov)	15-min	0.22 (2.43)**	−0.45 (−4.91)***	0.48 (16.80)***
	30-min	0.19 (2.30)**	−0.45 (−5.03)***	0.47 (14.50)***
	60-min	0.17 (1.59)	−0.42 (−3.11)***	0.42 (8.39)***

Table 7. Root Mean Squared Forecast Error

This table reports the daily out-of-sample RMSFEs ($\times 10^{-5}$) for all models relative to the true volatility proxies: 5-minute realized volatility (Panel A), median-based volatility (Panel B), and range-based volatility (Panel C). The out-of-sample period for each commodity futures contract is reported in Table 1.

Interval	Model	Aluminum	Copper	Fuel Oil	Sugar (Jan)	Sugar (Nov)
			Panel A: 5-Minute Volatility			
15-min	ARFIMA	5.065	18.767	23.295	6.372	6.699
	GARCH	6.343	27.982	33.415	10.236	9.369
	FIGARCH	5.866	24.223	32.233	8.604	9.697
	IGARCH	5.556	25.056	36.064	8.559	8.887
30-min	ARFIMA	5.072	18.819	23.474	6.467	6.796
	GARCH	6.855	23.381	28.183	10.061	8.930
	FIGARCH	6.117	21.988	21.916	8.626	8.922
	IGARCH	5.956	22.270	27.298	8.917	9.301
60-min	ARFIMA	5.078	18.981	23.507	6.509	6.944
	GARCH	6.845	21.939	25.094	8.737	8.370
	FIGARCH	5.764	20.788	22.649	7.993	8.597
	IGARCH	5.848	21.505	24.985	8.418	8.749
Daily	GARCH	5.081	18.912	23.474	7.315	6.728
	FIGARCH	5.052	18.606	23.476	7.101	6.728
	IGARCH	5.050	19.038	23.465	7.394	6.765
			Panel B: Median-Based Volatility			
15-min	ARFIMA	1.366	6.461	11.154	2.869	10.962
	GARCH	5.064	30.450	30.190	11.128	14.441
	FIGARCH	4.002	25.299	30.063	8.934	14.615
	IGARCH	3.687	26.740	33.895	8.927	13.893
30-min	ARFIMA	1.330	6.282	11.120	2.629	10.877
	GARCH	5.655	25.723	24.134	10.997	14.318
	FIGARCH	4.410	22.722	12.787	9.009	14.220
	IGARCH	4.258	23.808	22.800	9.518	14.655
60-min	ARFIMA	1.333	5.938	11.065	2.510	10.827
	GARCH	5.593	22.637	19.369	9.483	13.872
	FIGARCH	3.754	20.180	13.334	8.169	13.901
	IGARCH	4.013	21.828	19.164	9.033	14.196
Daily	GARCH	2.109	14.708	11.349	7.328	12.470
	FIGARCH	1.699	13.537	11.339	6.943	12.296
	IGARCH	1.962	15.196	11.337	7.471	12.516
			Panel C: Range-Based Volatility			
15-min	ARFIMA	1.994	5.581	14.147	5.466	5.464
	GARCH	5.190	26.133	35.255	12.659	10.395
	FIGARCH	4.178	20.592	36.589	10.617	10.641
	IGARCH	3.893	22.314	39.476	10.618	9.668
30-min	ARFIMA	1.963	5.685	14.106	5.299	5.314
	GARCH	5.747	21.612	24.335	12.516	10.217
	FIGARCH	4.571	19.046	18.225	10.705	10.117
	IGARCH	4.414	19.771	23.077	11.135	10.647
60-min	ARFIMA	1.957	5.674	14.053	5.249	5.206
	GARCH	5.660	18.789	20.009	11.066	9.560
	FIGARCH	3.933	16.443	16.070	9.871	9.505
	IGARCH	4.159	18.019	19.822	10.639	9.975
Daily	GARCH	2.429	11.556	13.902	9.058	7.782
	FIGARCH	2.129	10.530	13.905	8.711	7.514
	IGARCH	2.319	11.993	13.898	9.180	7.848

Table 8. Diebold and Mariano (1995) and West (1996) Test Results

The table reports the test statistics of the Diebold and Mariano (1995) and West (1996) test based on the Andrews and Monahan (1992) estimator. The benchmark models are those with the lowest RMSFE in Table 7. The forecast errors are computed relative to 5-minute realized volatility (Panel A), median-based volatility (Panel B), and range-based volatility (Panel C) measures. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The out-of-sample period for each commodity futures contract is reported in Table 1.

Commodity	Benchmark vs.	ARFIMA			GARCH				FIGARCH				IGARCH			
		15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
Panel A: 5-Minute Volatility																
Aluminium	IGARCH Daily	−0.27	−0.36	−0.41	−3.51***	−5.25***	−7.46***	−3.18***	−3.42***	−4.23***	−6.19***	−0.06	−2.15**	−4.32***	−7.02***	
Copper	FIGARCH Daily	−0.18	−0.23	−0.37	−4.45***	−3.68***	−3.79***	−2.13**	−2.35***	−3.51***	−4.12***		−3.95***	−3.32***	−3.69***	−2.63***
Fuel Oil	FIGARCH 30-min	−0.74	−0.82	−0.85	−0.98	−3.62***	−3.32***	−0.83	−1.41		−0.66	−0.84	−1.04	−3.56***	−3.23***	−0.83
Sugar (Jan)	ARFIMA 15-min		−2.57***	−2.47***	−4.45***	−6.29***	−5.74***	−1.96**	−3.68***	−5.20***	−4.57***	−1.67*	−3.46***	−5.42***	−5.32***	−2.07**
Sugar (Nov)	ARFIMA 15-min		−2.88***	−3.13***	−1.90*	−2.30**	−2.06**	−0.07	−2.02**	−2.32**	−1.81*	−0.08	−1.71*	−2.43***	−2.18**	−0.17
Panel B: Median-Based Volatility																
Aluminum	ARFIMA 30-min	−2.72***		−0.19	−6.93***	−6.95***	−7.39***	−4.92***	−6.47***	−6.78***	−7.33***	−3.52***	−7.15***	−7.29***	−7.75***	−4.84***
Copper	ARFIMA 60-min	−2.94***	−1.47		−5.48***	−2.75***	−2.42***	−2.21**	−3.52***	−6.07***	−3.34***	−1.31	−4.94***	−2.87***	−2.49***	−2.02**
Fuel Oil	ARFIMA 60-min	−0.89	−1.63		−1.30	−4.41***	−6.63***	−2.52***	−1.73*	−0.65	−2.19**	−2.53***	−1.34	−4.56***	−6.64***	−2.49***
Sugar (Jan)	ARFIMA 60-min	−5.41***	−3.67***		−6.95***	−8.39***	−10.78***	−3.18***	−7.84***	−9.20***	−10.09***	−3.20***	−6.90***	−8.61***	−11.62***	−3.27***
Sugar (Nov)	ARFIMA 60-min	−1.72*	−1.03		−4.14***	−5.05***	−5.44***	−3.17***	−3.93***	−5.13***	−4.55***	−3.21***	−3.87***	−4.95***	−5.18***	−3.20***
Panel C: Range-Based Volatility																
Aluminum	ARFIMA 60-min	−2.40***	−0.52		−7.32***	−7.08***	−7.86***	−3.51***	−6.77***	−6.91***	−7.62***	−1.79*	−7.67***	−7.56***	−7.99***	−3.03***
Copper	ARFIMA 15-min		−1.63	−0.77	−7.83***	−5.84***	−5.07***	−3.47***	−4.57***	−7.22***	−5.68***	−2.62***	−8.19***	−6.17***	−5.23***	−3.25***
Fuel Oil	IGARCH Daily	−1.30	−1.22	−1.06	−0.98	−3.83***	−4.48***	−0.69	−1.47	−2.58***	−3.44***	−0.61	−1.11	−3.90***	−4.46***	
Sugar (Jan)	ARFIMA 60-min	−3.45***	−1.19		−7.02***	−8.61***	−10.39***	−4.15***	−7.68***	−9.04***	−9.27***	−3.96***	−6.97***	−8.75***	−10.78***	−4.25***
Sugar (Nov)	ARFIMA 60-min	−3.25***	−2.33***		−4.04***	−5.43***	−6.40***	−3.30***	−4.23***	−5.85***	−5.93***	−3.44***	−3.73***	−5.34***	−6.12***	−3.33***

Table 9. Superior Predictive Ability Test Results

The table reports the Hansen (2005) SPA test results based on the MSFE. The benchmark models are those with the lowest RMSFE in Table 7. The forecast errors are computed relative to 5-minute realized volatility (Panel A), median-based volatility (Panel B), and range-based volatility (Panel C) measures. The null hypothesis is that the benchmark model is not inferior to the alternative models. The stationary bootstrap p-values are obtained using 10,000 replications. The out-of-sample period for each commodity futures contract is reported in Table 1.

Commodity	Benchmark	p-value
Panel A: 5-Minute Volatility		
Aluminum	IGARCH Daily	1.00
Copper	FIGARCH Daily	1.00
Fuel Oil	FIGARCH 30-min	0.99
Sugar (Jan)	ARFIMA 15-min	1.00
Sugar (Nov)	ARFIMA 15-min	1.00
Panel B: Median-Based Volatility		
Aluminum	ARFIMA 30-min	0.99
Copper	ARFIMA 60-min	1.00
Fuel Oil	ARFIMA 60-min	0.99
Sugar (Jan)	ARFIMA 60-min	1.00
Sugar (Nov)	ARFIMA 60-min	1.00
Panel C: Range-Based Volatility		
Aluminum	ARFIMA 60-min	0.99
Copper	ARFIMA 15-min	1.00
Fuel Oil	IGARCH Daily	1.00
Sugar (Jan)	ARFIMA 60-min	1.00
Sugar (Nov)	ARFIMA 60-min	1.00

Appendix

Table A1. Diebold and Mariano (1995) and West (1996) Test Results: 5-Minute Volatility Proxy

The table reports the test statistics of the Diebold and Mariano (1995) and West (1996) test based on the Andrews and Monahan (1992) estimator. Based on the results of the RMSFE presented in Table 7, the benchmark models are chosen in terms of increasing RMSFE. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The forecast errors for all models are computed relative to 5-minute measure of true volatility. The out-of-sample period for each commodity futures contract is reported in Table 1.

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
Panel A: Aluminium															
IGARCH Daily	−0.27	−0.36	−0.41	−3.51***	−5.25***	−7.46***	−3.18***	−3.42***	−4.23***	−6.19***	−0.06	−2.15**	−4.32***	−7.02***	
FIGARCH Daily	−0.45	−0.62	−0.67	−3.32***	−5.03***	−7.22***	−0.68	−3.14***	−3.99***	−5.89***		−1.93*	−3.93***	−6.29***	
ARFIMA 15-min		−0.81	−0.76	−3.23***	−4.94***	−6.99***	−0.23	−3.01***	−3.83***	−5.40***		−1.81*	−3.72***	−5.66***	
ARFIMA 30-min			−0.39	−3.21***	−4.91***	−6.97***	−0.13	−2.98***	−3.81***	−5.27***		−1.78*	−3.70***	−5.53***	
ARFIMA 60-min				−3.13***	−4.84***	−6.86***	−0.04	−2.87***	−3.70***	−5.06***		−1.70*	−3.55***	−5.28***	
GARCH Daily				−3.47***	−5.22***	−7.47***		−3.33***	−4.16***	−6.07***		−2.05**	−4.25***	−7.02***	
IGARCH 15-min				−5.06***	−5.68***	−4.17***		−2.60***	−3.61***	−0.92			−3.25***	−1.34	
FIGARCH 60-min				−1.77*	−3.76***	−7.49***		−0.49	−1.78*				−1.05	−1.92*	
IGARCH 60-min				−1.56	−3.61***	−7.57***		−0.09	−1.38				−0.63		
FIGARCH 15-min				−2.57***	−4.34***	−3.89***			−2.02**				−0.61		
IGARCH 30-min				−1.92*	−6.08***	−3.88***			−1.38						
FIGARCH 30-min				−1.11	−4.19***	−3.25***									
GARCH 15-min					−2.64***	−1.50									
GARCH 60-min					−0.04										
Panel B: Copper															
FIGARCH Daily	−0.18	−0.23	−0.37	−4.45***	−3.68***	−3.79***	−2.13**	−2.35***	−3.51***	−4.12***		−3.95***	−3.32***	−3.69***	−2.63***
ARFIMA 15-min		−1.14	−1.61	−3.94***	−2.31**	−1.92*	−0.16	−2.09**	−1.94*	−1.61		−3.30***	−1.92*	−1.74*	−0.28
ARFIMA 30-min			−1.48	−3.91***	−2.28**	−1.87*	−0.10	−2.07**	−1.89*	−1.53		−3.26***	−1.89*	−1.69*	−0.22
GARCH Daily			−0.07	−4.32***	−3.41***	−3.48***		−2.24**	−3.26***	−3.64***		−3.79***	−3.02***	−3.33***	−1.60
ARFIMA 60-min				−3.79***	−2.12**	−1.69*		−1.98**	−1.71*	−1.33		−3.09***	−1.73*	−1.51	−0.05
IGARCH Daily				−4.27***	−3.46***	−3.59***		−2.20**	−3.27***	−3.72***		−3.72***	−3.04***	−3.44***	
FIGARCH 60-min				−3.83***	−2.98***	−2.60***		−1.58	−2.16**			−3.04***	−2.16**	−2.04**	
IGARCH 60-min				−3.54***	−2.89***	−3.94***		−1.28	−0.89			−2.55***	−1.51		
GARCH 60-min				−3.33***	−2.41***			−1.08	−0.09			−2.22**	−0.67		
FIGARCH 30-min				−3.63***	−2.17**			−1.09				−2.62***	−0.59		
IGARCH 30-min				−3.68***	−4.64***			−0.99				−2.60***			
GARCH 30-min				−3.08***				−0.44				−1.57			
FIGARCH 15-min				−2.93***								−0.67			
IGARCH 15-min				−4.80***											

(Continued)

Table A1 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
								Panel C: Fuel Oil							
FIGARCH 30-min	−0.74	−0.82	−0.85	−0.98	−3.62***	−3.32***	−0.83	−1.41		−0.66	−0.84	−1.04	−3.56***	−3.23***	−0.83
FIGARCH 60-min	−0.73	−0.91	−0.97	−1.09	−2.52***	−3.04***	−0.95	−1.46			−0.96	−1.13	−2.50***	−2.98***	−0.94
ARFIMA 15-min		−2.57***	−2.38***	−1.08	−1.63	−1.02	−1.15	−1.48			−1.19	−1.16	−1.47	−0.98	−1.13
IGARCH Daily		−0.07	−0.35	−1.07	−1.55	−0.91	−1.02	−1.46			−0.38	−1.16	−1.39	−0.86	
ARFIMA 30-min			−0.95	−1.07	−1.55	−0.90	−0.01	−1.46			−0.02	−1.15	−1.39	−0.86	
GARCH Daily			−0.26	−1.07	−1.54	−0.90		−1.46			−0.06	−1.16	−1.38	−0.85	
FIGARCH Daily			−0.25	−1.07	−1.54	−0.90		−1.46				−1.16	−1.39	−0.86	
ARFIMA 60-min				−1.06	−1.54	−0.89		−1.45				−1.15	−1.38	−0.84	
IGARCH 60-min				−0.87	−1.83*	−2.57***		−1.02				−0.97	−1.62		
GARCH 60-min				−0.86	−1.79*			−1.01				−0.96	−1.57		
IGARCH 30-min				−0.64	−2.60***			−0.66				−0.77			
GARCH 30-min				−0.55				−0.53				−0.69			
FIGARCH 15-min				−0.24								−0.80			
GARCH 15-min												−1.29			
								Panel D: Sugar (Jan)							
ARFIMA 15-min		−2.57***	−2.47***	−4.45***	−6.29***	−5.74***	−1.96**	−3.68***	−5.20***	−4.57***	−1.67*	−3.46***	−5.42***	−5.32***	−2.07**
ARFIMA 30-min			−1.38	−4.35***	−6.08***	−5.29***	−1.73*	−3.52***	−4.86***	−4.10***	−1.39	−3.28***	−5.17***	−4.84***	−1.87*
ARFIMA 60-min				−4.31***	−6.01***	−5.15***	−1.64	−3.46***	−4.74***	−3.96***	−1.29	−3.22***	−5.08***	−4.70***	−1.78*
FIGARCH Daily				−3.76***	−5.72***	−4.77***	−2.20**	−2.58***	−3.59***	−2.07**		−2.27**	−4.11***	−3.67***	−2.88***
GARCH Daily				−3.59***	−5.34***	−3.74***		−2.22**	−3.00***	−1.43		−1.93*	−3.47***	−2.68***	−5.75***
IGARCH Daily				−3.50***	−5.17***	−3.46***		−2.07**	−2.78***	−1.24		−1.80*	−3.26***	−2.43***	
FIGARCH 60-min				−3.42***	−5.06***	−5.88***		−1.48	−2.82***			−1.35	−3.21***	−4.11***	
IGARCH 60-min				−2.70***	−3.93***	−5.86***		−0.42	−0.77			−0.31	−1.60		
IGARCH 15-min				−5.77***	−3.89***	−0.39		−0.15	−0.21				−1.07		
FIGARCH 15-min				−3.51***	−2.96***	−0.31			−0.07				−0.72		
FIGARCH 30-min				−3.07***	−5.04***	−0.41							−1.55		
GARCH 60-min				−2.27**	−3.28***								−0.58		
IGARCH 30-min				−2.59***	−7.43***										
30-min GARCH				−0.39											

(Continued)

Table A1 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
							Panel E: Sugar (Nov)								
ARFIMA 15-min		−2.88***	−3.13***	−1.90*	−2.30**	−2.06**	−0.07	−2.02**	−2.32**	−1.81*	−0.08	−1.71*	−2.43***	−2.18**	−0.17
GARCH Daily		−0.17	−0.52	−1.82*	−2.16**	−1.83*		−1.98**	−2.31**	−1.75*	−0.01	−1.64	−2.29**	−1.93*	−3.21***
FIGARCH Daily		−0.19	−0.58	−1.83*	−2.16**	−1.83*		−1.98**	−2.31**	−1.75*		−1.65*	−2.28**	−1.92*	−0.47
IGARCH Daily		−0.07	−0.42	−1.80*	−2.13**	−1.78*		−1.95*	−2.27**	−1.71*		−1.61	−2.26**	−1.89*	
ARFIMA 30-min			−2.77***	−1.83*	−2.19**	−1.93*		−1.95*	−2.22**	−1.71*		−1.63	−2.33***	−2.07**	
ARFIMA 60-min				−1.73*	−2.05**	−1.76*		−1.86*	−2.06**	−1.58		−1.53	−2.22**	−1.93*	
GARCH 60-min				−1.67*	−2.23**			−1.64	−1.74*	−0.87		−1.06	−2.79***	−2.69***	
FIGARCH 60-min				−1.52	−1.00			−1.77*	−1.29			−0.66	−1.95*	−0.73	
IGARCH 60-min				−1.24	−0.90			−1.31	−0.63			−0.34	−2.17**		
IGARCH 15-min				−3.24***	−0.13			−1.57	−0.08				−1.44		
FIGARCH 30-min				−0.79	−0.03			−1.31					−1.35		
GARCH 30-min				−0.98				−1.09					−3.19***		
IGARCH 30-min				−0.18				−0.61							
GARCH 15-min								−0.72							

Table A2. Diebold and Mariano (1995) and West (1996) Test Results: Median-Based Volatility Proxy

The table reports the test statistics of the Diebold and Mariano (1995) and West (1996) test based on the Andrews and Monahan (1992) estimator. Based on the results of the RMSFE presented in Table 7, the benchmark models are chosen in terms of increasing RMSFE. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The forecast errors for all models are computed relative to the median-based measure of true volatility. The out-of-sample period for each commodity futures contract is reported in Table 1.

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
							Panel A: Aluminum								
ARFIMA 30-min	−2.72***		−0.19	−6.93***	−6.95***	−7.39***	−4.92***	−6.47***	−6.78***	−7.33***	−3.52***	−7.15***	−7.29***	−7.75***	−4.84***
ARFIMA 60-min	−1.51			−6.96***	−7.01***	−7.49***	−5.21***	−6.49***	−6.79***	−7.38***	−3.70***	−7.19***	−7.39***	−7.89***	−5.15***
ARFIMA 15-min				−6.90***	−6.91***	−7.36***	−4.67***	−6.40***	−6.73***	−7.31***	−3.26***	−7.05***	−7.22***	−7.71***	−4.55***
FIGARCH Daily				−6.65***	−6.75***	−7.31***	−6.19***	−6.03***	−6.44***	−7.12***		−6.49***	−6.91***	−7.60***	−6.46***
IGARCH Daily				−6.38***	−6.56***	−7.15***	−4.48***	−5.65***	−6.11***	−6.62***		−5.93***	−6.53***	−7.19***	
GARCH Daily				−6.24***	−6.47***	−7.12***		−5.40***	−5.92***	−6.39***		−5.57***	−6.32***	−7.05***	
IGARCH 15-min				−6.62***	−5.66***	−4.77***		−1.81*	−3.30***	−0.31			−3.35***	−1.38	
FIGARCH 60-min				−4.03***	−4.96***	−6.32***		−1.00	−2.79***				−2.25**	−3.98***	
FIGARCH 15-min				−4.68***	−4.87***	−4.40***			−2.28**				−1.16	−0.04	
IGARCH 60-min				−3.27***	−4.50***	−6.92***			−1.61				−1.08		
IGARCH 30-min				−3.60***	−6.46***	−4.11***			−0.89						
FIGARCH 30-min				−2.81***	−4.94***	−3.73***									
GARCH 15-min					−2.57***	−1.55									
GARCH 60-min					−0.19										
							Panel B: Copper								
ARFIMA 60-min	−2.94***	−1.47		−5.48***	−2.75***	−2.42***	−2.21**	−3.52***	−6.07***	−3.34***	−1.31	−4.94***	−2.87***	−2.49***	−2.02**
ARFIMA 30-min	−2.81***			−5.48***	−2.73***	−2.43***	−2.16**	−3.50***	−6.22***	−3.38***	−1.28	−4.95***	−2.86***	−2.51***	−1.98**
ARFIMA 15-min				−5.49***	−2.76***	−2.43***	−2.00**	−3.50***	−6.27***	−3.35***	−1.29	−4.96***	−2.91***	−2.51***	−1.84*
FIGARCH Daily				−4.95***	−3.26***	−3.47***	−2.82***	−2.80***	−6.97***	−4.60***		−4.32***	−3.53***	−3.75***	−2.87***
GARCH Daily				−4.73***	−3.08***	−2.89***		−2.60***	−6.51***	−3.83***		−4.05***	−3.32***	−3.09***	−0.89
IGARCH Daily				−4.65***	−3.18***	−3.09***		−2.52***	−6.52***	−3.87***		−3.95***	−3.41***	−3.33***	
FIGARCH 60-min				−4.00***	−3.86***	−3.02***		−1.53	−4.93***			−3.06***	−3.93***	−3.10***	
IGARCH 60-min				−3.46***	−3.19***	−1.72*		−1.09	−1.58			−2.34***	−2.47***		
GARCH 60-min				−3.20***	−2.96***			−0.85	−0.14			−1.98**	−1.61		
FIGARCH 30-min				−3.33***	−3.33***			−0.84				−2.12**	−1.90*		
IGARCH 30-min				−3.25***	−2.03**			−0.53				−1.80*			
FIGARCH 15-min				−3.43***	−0.16							−0.89			
GARCH 30-min				−2.44***								−0.64			
IGARCH 15-min				−5.82***											

(Continued)

Table A2 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
							Panel C: Fuel Oil								
ARFIMA 60-min	−0.89	−1.63		−1.30	−4.41***	−6.63***	−2.52***	−1.73*	−0.65	−2.19**	−2.53***	−1.34	−4.56***	−6.64***	−2.49***
ARFIMA 30-min	−0.46			−1.30	−4.39***	−6.59***	−1.76*	−1.72*	−0.63	−2.16**	−1.76*	−1.33	−4.53***	−6.60***	−1.72*
ARFIMA 15-min				−1.29	−4.36***	−6.60***	−1.06	−1.72*	−0.63	−2.22***	−1.04	−1.33	−4.50***	−6.61***	−1.02
IGARCH Daily				−1.30	−4.39***	−6.44***	−1.75*	−1.72*	−0.54	−1.83*	−0.08	−1.33	−4.53***	−6.43***	
FIGARCH Daily				−1.30	−4.38***	−6.44***	−0.46	−1.72*	−0.54	−1.83*		−1.33	−4.53***	−6.43***	
GARCH Daily				−1.30	−4.38***	−6.42***		−1.72*	−0.54	−1.82*		−1.33	−4.52***	−6.41***	
FIGARCH 30-min				−1.03	−4.03***	−3.94***		−1.51		−0.34		−1.15	−4.02***	−3.83***	
FIGARCH 60-min				−1.15	−3.53***	−4.49***		−1.57				−1.22	−3.49***	−4.53***	
IGARCH 60-min				−0.84	−2.50***	−4.68***		−1.13				−0.97	−2.17**		
GARCH 60-min				−0.83	−2.42***			−1.12				−0.96	−2.07**		
IGARCH 30-min				−0.59	−3.67***			−0.79				−0.77			
GARCH 30-min				−0.49				−0.66				−0.68			
FIGARCH 15-min				−0.02								−0.73			
GARCH 15-min												−1.28			
							Panel D: Sugar (Jan)								
ARFIMA 60-min	−5.41***	−3.67***		−6.95***	−8.39***	−10.78***	−3.18***	−7.84***	−9.20***	−10.09***	−3.20***	−6.90***	−8.61***	−11.62***	−3.27***
ARFIMA 30-min	−4.72***			−6.91***	−8.33***	−10.73***	−3.15***	−7.78***	−9.14***	−10.02***	−3.15***	−6.84***	−8.54***	−11.53***	−3.24***
ARFIMA 15-min				−6.86***	−8.26***	−10.64***	−3.03***	−7.66***	−9.01***	−9.87***	−2.99***	−6.76***	−8.44***	−11.50***	−3.11***
FIGARCH Daily				−4.60***	−6.25***	−4.64***	−3.33***	−3.12***	−3.32***	−1.48		−2.61***	−4.24***	−3.33***	−4.25***
GARCH Daily				−4.33***	−5.72***	−3.73***		−2.57***	−2.70***	−0.99		−2.13**	−3.52***	−2.50***	−6.14***
IGARCH Daily				−4.19***	−5.50***	−3.37***		−2.33***	−2.43***	−0.79		−1.93*	−3.26***	−2.22**	
FIGARCH 60-min				−4.29***	−5.60***	−6.54***		−1.85*	−3.10***			−1.69*	−3.80***	−6.04***	
IGARCH 15-min				−6.87***	−5.18***	−1.15		−0.03	−0.26				−1.83*	−0.22	
FIGARCH 15-min				−4.73***	−4.75***	−1.30			−0.26				−1.54	−0.23	
FIGARCH 30-min				−3.92***	−5.93***	−1.48							−2.68***	−0.07	
IGARCH 60-min				−2.99***	−4.01***	−4.81***							−1.37		
GARCH 60-min				−2.43***	−3.33***								−0.10		
IGARCH 30-min				−3.20***	−7.39***										
GARCH 30-min				−0.31											

(Continued)

Table A2 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
Panel E: Sugar (Nov)															
ARFIMA 60-min	−1.72*	−1.03		−4.14***	−5.05***	−5.44***	−3.17***	−3.93***	−5.13***	−4.55***	−3.21***	−3.87***	−4.95***	−5.18***	−3.20***
ARFIMA 30-min	−2.53***			−4.10***	−5.05***	−5.48***	−3.08***	−3.91***	−5.16***	−4.47***	−3.20***	−3.82***	−4.93***	−5.16***	−3.11***
ARFIMA 15-min				−4.02***	−4.97***	−5.40***	−2.87***	−3.86***	−5.11***	−4.34***	−3.06***	−3.73***	−4.85***	−5.05***	−2.90***
FIGARCH Daily				−2.39***	−3.25***	−3.08***	−3.10***	−2.41***	−3.13***	−2.30**		−1.92*	−3.45***	−3.23***	−3.64***
GARCH Daily				−2.16**	−2.92***	−2.62***		−2.22**	−2.82***	−1.97**		−1.68*	−3.19***	−2.85***	−3.71***
IGARCH Daily				−2.11**	−2.84***	−2.51***		−2.16**	−2.74***	−1.89*		−1.61	−3.11***	−2.76***	
GARCH 60-min				−1.31	−2.29**			−1.33	−1.49	−0.18		−0.06	−3.08***	−2.74***	
IGARCH 15-min				−6.17***	−1.47			−2.12**	−1.05	−0.02			−2.61***	−0.90	
FIGARCH 60-min				−1.47	−1.80*			−1.59	−1.68*				−2.83***	−2.05**	
IGARCH 60-min				−0.64	−0.76			−0.81	−0.11				−2.29**		
FIGARCH 30-min				−0.59	−0.54			−1.00					−2.02**		
GARCH 30-min				−0.37				−0.62					−3.70***		
GARCH 15-min								−0.57					−0.69		
FIGARCH 15-min													−0.09		

Table A3. Diebold and Mariano (1995) and West (1996) Test Results: Range-Based Volatility Proxy

The table reports the test statistics of the Diebold and Mariano (1995) and West (1996) test based on the Andrews and Monahan (1992) estimator. Based on the results of the RMSFE presented in Table 7, the benchmark models are chosen in terms of increasing RMSFE. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The forecast errors for all models are computed relative to the range-based measure of true volatility. The out-of-sample period for each commodity futures contract is reported in Table 1.

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
							Panel A: Aluminum								
ARFIMA 60-min	−2.40***	−0.52		−7.32***	−7.08***	−7.86***	−3.51***	−6.77***	−6.91***	−7.62***	−1.79*	−7.67***	−7.56***	−7.99***	−3.03***
ARFIMA 30-min	−3.62***			−7.29***	−7.03***	−7.83***	−3.52***	−6.75***	−6.90***	−7.61***	−1.80*	−7.64***	−7.49***	−7.95***	−3.05***
ARFIMA 15-min				−7.25***	−6.99***	−7.82***	−3.34***	−6.67***	−6.85***	−7.57***	−1.49	−7.55***	−7.42***	−7.93***	−2.82***
FIGARCH Daily				−6.94***	−6.85***	−7.77***	−5.61***	−6.30***	−6.60***	−7.37***		−6.91***	−7.01***	−8.00***	−5.20***
IGARCH Daily				−6.63***	−6.63***	−7.56***	−4.98***	−5.85***	−6.24***	−6.79***		−6.23***	−6.60***	−7.54***	
GARCH Daily				−6.48***	−6.54***	−7.50***		−5.58***	−6.03***	−6.52***		−5.84***	−6.39***	−7.36***	
IGARCH 15-min				−6.82***	−5.63***	−5.23***		−1.76*	−3.26***	−0.21			−3.32***	−1.32	
FIGARCH 60-min				−4.26***	−4.97***	−6.79***		−1.09	−2.98***				−2.44***	−3.27***	
IGARCH 60-min				−3.51***	−4.58***	−7.41***		−0.08	−1.78*				−1.28		
FIGARCH 15-min				−4.72***	−4.84***	−4.48***			−2.33***				−1.14		
IGARCH 30-min				−3.74***	−6.54***	−4.50***			−0.94						
FIGARCH 30-min				−2.80***	−4.78***	−3.80***									
GARCH 15-min					−2.55***	−1.52									
GARCH 60-min					−0.30										
							Panel B: Copper								
ARFIMA 15-min		−1.63	−0.77	−7.83***	−5.84***	−5.07***	−3.47***	−4.57***	−7.22***	−5.68***	−2.62***	−8.19***	−6.17***	−5.23***	−3.25***
ARFIMA 60-min		−0.11		−7.89***	−5.88***	−5.14***	−3.60***	−4.60***	−7.17***	−5.76***	−2.76***	−8.30***	−6.21***	−5.30***	−3.39***
ARFIMA 30-min				−7.85***	−5.89***	−5.13***	−3.58***	−4.58***	−7.23***	−5.77***	−2.72***	−8.24***	−6.22***	−5.30***	−3.36***
FIGARCH Daily				−6.71***	−6.12***	−6.02***	−2.72***	−3.60***	−7.47***	−6.95***		−6.47***	−6.56***	−6.41***	−2.76***
GARCH Daily				−6.46***	−5.72***	−5.49***		−3.36***	−6.92***	−5.93***		−6.11***	−6.01***	−5.79***	−1.73*
IGARCH Daily				−6.30***	−5.77***	−5.71***		−3.23***	−6.83***	−5.87***		−5.86***	−6.03***	−6.05***	
FIGARCH 60-min				−5.02***	−4.83***	−3.49***		−1.80*	−4.81***			−4.04***	−4.61***	−3.24***	
IGARCH 60-min				−4.20***	−4.13***	−3.28***		−1.15	−1.72*			−2.78***	−2.75***		
GARCH 60-min				−3.80***	−3.62***			−0.81	−0.41			−2.23**	−1.54		
FIGARCH 30-min				−4.04***	−3.42***			−0.72				−2.50***	−1.42		
IGARCH 30-min				−3.72***	−4.04***			−0.40				−1.95*			
FIGARCH 15-min				−4.16***	−0.51							−1.35			
GARCH 30-min				−2.65***								−0.50			
IGARCH 15-min				−5.87***											

(Continued)

Table A3 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
								Panel C: Fuel Oil							
IGARCH Daily	−1.30	−1.22	−1.06	−0.98	−3.83***	−4.48***	−0.69	−1.47	−2.58***	−3.44***	−0.61	−1.11	−3.90***	−4.46***	
GARCH Daily	−1.27	−1.19	−1.02	−0.98	−3.83***	−4.47***		−1.47	−2.59***	−3.43***	−0.25	−1.11	−3.90***	−4.45***	
FIGARCH Daily	−1.24	−1.15	−0.98	−0.98	−3.83***	−4.47***		−1.47	−2.57***	−3.41***		−1.11	−3.90***	−4.45***	
ARFIMA 60-min	−0.74	−0.85		−0.98	−3.76***	−4.22***		−1.47	−2.52***	−3.13***		−1.11	−3.81***	−4.19***	
ARFIMA 30-min	−0.53			−0.98	−3.75***	−4.17***		−1.47	−2.52***	−3.11***		−1.11	−3.79***	−4.14***	
ARFIMA 15-min				−0.98	−3.74***	−4.16***		−1.47	−2.50***	−3.25***		−1.11	−3.78***	−4.13***	
FIGARCH 60-min				−0.92	−3.10***	−2.80***		−1.39	−1.47			−1.06	−3.02***	−2.74***	
FIGARCH 30-min				−0.91	−1.69*	−0.66		−1.37				−1.05	−1.42	−0.60	
IGARCH 60-min				−0.77	−2.43***	−3.70***		−1.17				−0.92	−2.10**		
GARCH 60-min				−0.76	−2.36***			−1.16				−0.91	−2.01**		
IGARCH 30-min				−0.63	−3.54***			−0.98				−0.80			
GARCH 30-min				−0.57				−0.91				−0.75			
GARCH 15-min								−0.28				−1.37			
FIGARCH 15-min												−0.72			
								Panel D: Sugar (Jan)							
ARFIMA 60-min	−3.45***	−1.19		−7.02***	−8.61***	−10.39***	−4.15***	−7.68***	−9.04***	−9.27***	−3.96***	−6.97***	−8.75***	−10.78***	−4.25***
ARFIMA 30-min	−3.95***			−7.03***	−8.63***	−10.52***	−4.15***	−7.74***	−9.16***	−9.39***	−3.96***	−6.99***	−8.80***	−10.93***	−4.25***
ARFIMA 15-min				−7.02***	−8.60***	−10.55***	−4.04***	−7.74***	−9.12***	−9.34***	−3.83***	−6.96***	−8.78***	−10.98***	−4.14***
FIGARCH Daily				−4.76***	−6.42***	−5.54***	−3.62***	−3.34***	−3.62***	−1.81*		−2.80***	−4.53***	−4.05***	−4.47***
GARCH Daily				−4.47***	−5.90***	−4.25***		−2.76***	−2.96***	−1.13		−2.31**	−3.78***	−2.91***	−6.88***
IGARCH Daily				−4.33***	−5.69***	−3.89***		−2.54***	−2.70***	−0.93		−2.12**	−3.52***	−2.60***	
FIGARCH 60-min				−4.41***	−5.81***	−7.98***		−2.00**	−3.28***			−1.85*	−3.97***	−6.73***	
FIGARCH 15-min				−4.90***	−4.77***	−1.16			−0.34			−0.01	−1.51	−0.06	
IGARCH 15-min				−6.94***	−5.26***	−0.99			−0.29				−1.73*	−0.05	
IGARCH 60-min				−3.09***	−4.08***	−5.68***			−0.21				−1.50		
FIGARCH 30-min				−3.93***	−5.91***	−1.17							−2.39***		
GARCH 60-min				−2.52***	−3.41***								−0.22		
IGARCH 30-min				−3.26***	−7.71***										
GARCH 30-min				−0.36											

(Continued)

Table A3 – *Continued*

→ Competing Model	ARFIMA			GARCH				FIGARCH				IGARCH			
↓ Benchmark Model	15-min	30-min	60-min	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily	15-min	30-min	60-min	Daily
							Panel E: Sugar (Nov)								
ARFIMA 60-min	−3.25***	−2.33***		−4.04***	−5.43***	−6.40***	−3.30***	−4.23***	−5.85***	−5.93***	−3.44***	−3.73***	−5.34***	−6.12***	−3.33***
ARFIMA 30-min	−3.36***			−3.97***	−5.35***	−6.34***	−3.18***	−4.15***	−5.74***	−5.85***	−3.34***	−3.65***	−5.27***	−6.06***	−3.22***
ARFIMA 15-min				−3.91***	−5.29***	−6.29***	−3.00***	−4.09***	−5.66***	−5.76***	−3.15***	−3.58***	−5.21***	−6.01***	−3.04***
FIGARCH Daily				−2.37***	−3.39***	−3.51***	−3.17***	−2.58***	−3.45***	−2.85***		−1.88*	−3.62***	−3.83***	−3.67***
GARCH Daily				−2.18**	−3.13***	−3.14***		−2.40***	−3.14***	−2.54***		−1.67*	−3.41***	−3.54***	−4.22***
IGARCH Daily				−2.13**	−3.04***	−3.00***		−2.34***	−3.05***	−2.43***		−1.61	−3.33***	−3.42***	
FIGARCH 60-min				−1.40	−2.19**	−0.34		−1.66*	−2.10**			−0.29	−2.86***	−2.89***	
GARCH 60-min				−1.15	−1.90*			−1.40	−1.58			−0.17	−2.54***	−3.46***	
IGARCH 15-min				−6.36***	−1.41			−2.50***	−1.06				−2.62***	−0.53	
IGARCH 60-min				−0.64	−0.88			−0.93	−0.44				−1.96*		
FIGARCH 30-min				−0.51	−0.45			−1.05					−1.99**		
GARCH 30-min				−0.39				−0.73					−4.20***		
GARCH 15-min								−0.71					−0.60		
FIGARCH 15-min													−0.01		